Genetic and ecological investigations in the thistle genus *Onopordum* with special emphases on hybridization and evolution

Peter O'Hanlon

A thesis submitted for the degree of Doctor of Philosophy
Australian National University

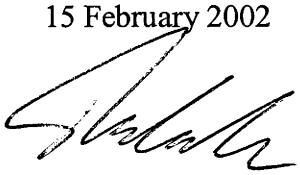
Library copy
February 2002

Declaration

I certify that the substance of this thesis has not already been submitted for any degree and is not currently submitted for any other degree.

I certify that the help received in preparing this thesis, and all sources used, have been acknowledged in this thesis.

Peter O'Hanlon
15 February 2002



Acknowledgements

This project was funded by the Department of Employment, Education, Training and Youth Affairs, the Australian Research Council and the Cooperative Research Centre for Weed Management Systems through an Australian Postgraduate Award (Industry) scholarship. Additional funding from the CRC for Weed Management Systems and the French Government Scientific and Cultural Exchange Program administered by the French Embassy, Canberra was also appreciated.

I would like to thank the following people for their support, advice and guidance;

- Dr. Rod Peakall (Chair of supervisory panel and primary supervisor) for sharing his expertise, endless enthusiasm and knowledge with me, and for providing me with a rounded academic training.
- Dr. David Briese (Co-supervisor) for providing the initial inspiration for this project, continual direction and assessment of my research in good humour and always being able to set my work within a context of Weed Science.
- My two examiners who provided helpful final suggestions to this thesis.
- My friends in BoZo and CSIRO: Occasionally you've made my life pretty normal.
- My advisors Dr. Siegy Krauss and Dr. Mike Crisp for guidance.
- A special thanks to Dr. Mike Ramsey, Dr. Richard Groves, Prof. Rick Roush and John Kerin for special encouragement and interest in my work.
- The technical assistance from the workshop (Bruce and Allan), in photography (Keith) and in the laboratory (Simon, Wes, Jeanette, Ginny, Ian)
- To my overseas friends and colleagues who added an extra dimension to my training and to this thesis; especially Stefan Bottcher, Isabelle Olivieri, John Scott, Eva Haeffner, Peter Hein (who also provided information on the distributions of *Onopordum*), Torsten, Thierry, and Jose.
- To those who have read much of my work and provided editorial input; Mike Ramsey, Siegy Krauss and especially Rod Peakall and David Briese.
- Seed for several analyses were provided by Thierry Thomann and Jose Serin (CSIRO European Laboratory, Montpellier, France), Eva Haeffner (Botanisches Museum and Botanische Garten, Berlin, Germany) and W. Richter (Botanisches Garten der Georg August Universitat, Gottingen, Germany) and Isabelle Olivieri (University of Montpellier, France).
- Personal thanks my family who have always been more interested in things other than my research.

Table of Contents

Preface

I have presented this thesis as a series of ten chapters. I have attempted to make each chapter discrete to facilitate the reading of individual chapters without unnecessary reference to other chapters, and consequently each has its own introduction and conclusions. However, they form a continuous thread from front the first to the last chapter and reading them in this order will provide the reader with the most benefit. Several of these chapters have been written based on papers that have previously been published. These publications have been included as appendicies. Each of these chapters use these publications only as a basis, and additional observations and interpretations have been included here. An introductory chapter to define the scope and aims of the thesis, and a final concluding chapter integrating the findings are also provided. As each chapter stands alone, some repetition is inevitable. To minimise this repetition, especially in relation to the sections relating to materials and methods, some cross-referencing is used. However, adequate detail is provided which should avoid the need to follow each of the cross-references.

Chapter One *Abstract*

Most organisms can be studied at both applied and theoretical levels. Pest species including weedy plants are one group of organisms that are particularly amenable to such dual investigation. The inspiration for this thesis flowed from the practical tasks of weed management, particularly the biological control of *Onopordum* thistles. However, as the study progressed, it became apparent that the system could be extremely valuable for observing evolutionary patterns and drawing theoretical conclusions, specifically relevant to the consequences of hybridization between invading taxa.

Genetic and ecological studies were performed in the genus *Onopordum* (*Asteraceae*), in both the Australian part of their alien distribution, and in France, part of their native range. To place the genus *Onopordum* within an evolutionary context, a phylogenetic analysis of the Carduineae (Cardueae; *Asteraceae*) was conducted. Chloroplast DNA sequence data from two regions (the intron between *psb*A-*trn*H, and the spacer and intron between *trn*L-*trn*F) revealed low levels of variation, with much of the variation being unique to individual genera. A PCR-based test for homology among AFLP fragments was developed, revealing that size homoplasy prevented phylogenetic analysis of AFLP data at higher, but not at lower taxonomic levels. A combination of cpDNA sequence, AFLP, and morphological data revealed that *Onopordum* was most closely related to two genera *Ptilostemon* and *Cynara* (Artichoke). As AFLPs were shown to have low frequencies of homoplasy within the genus *Onopordum* and *Cynara*, they were used to obtain genetic data for estimating interrelationships between European species of *Onopordum*, and to reconstruct a biogeographic history. As has been found with several other European taxa, a strong divide between representatives

from the Iberian peninsula (Spain) and the Aegean region (Greece, Turkey and Balkans) was observed, with an additional group containing the only species present in the northern parts of the genus' distribution. This pattern contrasted strongly with morphological variation which showed no geographic pattern, with most variants being present throughout the distribution. Such a pattern is consistent with isolation into southern refugia and post-glacial recolonization. It is possible that the range of extant morphological variants were conserved *via* hybridization in restricted refugia, and subsequently sorted by selection during recolonization. Simulation studies show that such a scenario may be most common when selection coefficients in different habitats for parental and hybrid species are intermediate, when there are large areas of suitable habitat in which hybrids are selected for, and when environmental change is relatively gradual.

Morphological and genetic analysis of *Onopordum* in Australia revealed a hybrid complex between *O. acanthium* (Scotch or cotton thistle) and *O. illyricum* (Illyrian thistle), with hybrids occupying much of the Australian range. Further detailed within population genetic analysis revealed that populations consisted of a relatively narrow subset of the entire hybrid complex's variation, indicative of low founder numbers, often by hybrid genotypes. In addition, nearly all populations analyzed were strongly divergent from one-another. Spatial autocorrelation analysis of morphological variation revealed strong local structure, consistent with populations being founded through immigration from neighboring sites (usually within 60km), although occasional long-distance dispersal was also identified. During the invasion of the novel habitat in Australia, hybrids could have become isolated from parental genotypes, without

biologically determined reproductive isolation or fitness advantages, through founder cascade, allowing allopatric hybrid evolution.

Analysis of bioclimatic correlations with species' distributions in Australia indicated that favorable habitats for *O. acanthium* and *O. illyricum* were mostly distinct, with some overlapping less favourable habitat. Hybrid populations were most abundant in such regions, supporting the concept of a widespread 'hybrid habitat' in Australia that occurs only rarely in Europe. Since the invasion of Australia, hybrids may have been released from selection pressures that maintained narrow hybrid zones in Europe, allowing the displacement of parents by hybrids in certain contexts.

Several early life-history fitness components were examined to compare relative advantages of parental and hybrid species. *O. acanthium* and *O. illyricum* differ in many early life-history traits such as seed mass, seed number, time to germination, and rate of seedling growth. Hybrids between these species were found to be variable though generally intermediate in these traits, thus providing opportunities for segregation of these characters, and adaptation to slightly altered or intermediate habitats. Though hybrid seed was generally less viable, total seed biomass was similar to that of parents. Competitive growth experiments in the glasshouse revealed similar patterns; that hybrids were more variable in response, but they were not inherently inferior to parents.

Since hybrids can be as fit as their parents, and may show a preference for habitats in Australia that are rare in Europe, the outcome of hybridization for species boundaries may be environmentally determined. Hybridization has been an important factor

shaping the refugial isolation and recolonization of Europe, as well as the contemporary patterns of invasion and adaptation in Australian *Onopordum*.

Chapter Two *A review of new PCR-based genetic markers and their utility to weed ecology*[1]

Before commencing laboratory-based genetic research, it is critical to explore all of the technical options available. Several recent molecular developments provide new genetic tools for addressing difficult problems in weed ecology. In this review I describe some of the techniques and the DNA markers they generate (including AFLPs, SSRs or microsatellites, and intron-PCR), contrast their relative advantages and disadvantages, and discuss how they may be used in weed research. As these new markers generally reveal higher levels of variation than other techniques, they promise to improve our understanding of breeding systems, assist in determining the origin/s of invaders, help resolve taxonomic boundaries and relationships between closely related taxa and enable the identification of phenotypic linked markers. In addition, compared with other techniques, some of these markers may be more cost effective, less technically demanding and more reliable. Regardless of the marker adopted, all genetic studies will benefit from careful consideration of experimental design and the formulation of testable hypotheses with practical outcomes.

The advent of the Polymerase Chain Reaction (PCR; Saiki *et al.*, 1998) has enabled the development of powerful genetic markers for the measurement of genotypic variation. By measuring genotype, rather than phenotype, genetic markers avoid complicating environmental effects and provide ideal tools for assessing genetic variation, identifying species and other locally adapted forms, and defining genetic relationships. Nissen *et al.* (1995) explored the utility of two genetic markers, Randomly Amplified Polymorphic DNA or RAPDs (the amplification of many DNA fragments by short primers of arbitrary sequence) and Restriction Fragment Length Polymorphism (RFLPs) for investigating the ecological genetics of weeds. These tools have become well established and are currently used widely in weed research. Since that review, an additional suite of molecular markers have become widely available and it is my

[1]A substantial part of the content of this chapter is drawn from O'Hanlon, Peakall & Briese (2000)

purpose here to: (1) describe and evaluate these genetic methods; and (2) review some of the recent applications of molecular markers in weed ecology.

Recent developments in genetic marker technology

Since the review of Nissen *et al.* (1995), two major classes of PCR-based procedures have been significantly advanced; sequence-tagged-site PCR (STS-PCR), and a combination of STS-PCR and arbitrarily primed PCR (such as that utilised in the RAPDs). Here, I describe two types of genetic marker that typify these different approaches and discuss their advantages and limitations. It is important to recognise that there is no perfect genetic marker and newer techniques are not necessarily "better" than the older ones. The choice of an appropriate method will be dependent on the aims of the study, technical considerations, availability of appropriate laboratory facilities, and the cost of developing and applying the method (Morell *et al.*, 1995; Peakall, 1997). I have provided a qualitative summary of the characteristics of the various genetic markers for comparison in Table 2.1 which may assist new users to select the most appropriate marker for their particular needs. The following discussion will highlight the features of these different marker systems.

Sequence-Tagged-Site PCR

Arbitrarily primed methods (such as RAPDs) utilise non-specific primers to amplify anonymous parts of the genome flanked in between priming sites in opposite orientation. In contrast STS-PCR uses two different specific primers, complementary in DNA sequence to opposite strands of conserved organism DNA, to amplify the

intervening sequence. A desirable feature of STS-PCR is that it produces a codominant single-locus genetic marker, with both alleles being expressed in diploid organisms for each locus. Even though multiple alleles may be revealed in a population, for diploids, a maximum of two alleles can be amplified, and one can distinguish heterozygotes from homozygotes. Consequently, the DNA profiles are generally much easier to score than multi-locus profiles, characterised by many bands but with no capacity to distinguish alleles from loci (cf. Figs 2.1 & Fig. 2.2). Another advantage of STS-PCR is that a positive result is always achieved (except in the case of null alleles) from a successful reaction. In contrast, with a multi-locus method (see below), partial failure of the PCR reaction may produce false polymorphisms or band absences (Morel *et al.*, 1995, Peakall, 1997).

Microsatellites or Simple Sequence Repeats (SSRs) are one promising class of extremely variable genetic markers revealed by STS-PCR. They have been utilised for over a decade, but largely restricted to some well-known animals, and fewer plants (Peakall *et al.*, 1998). SSRs consist of tandemly repeated units of a short nucleotide motif, 1-6 bps long. Di-, tri- and tetra- nucleotide repeats are the most common (e.g. (CA)n, (AAT)n, (GATA)n, respectively), and are widely distributed throughout the genomes of plants and animals (Jarne & Lagoda, 1996). The hypervariability of SSR loci is a consequence of exceptionally high mutation rates for these nucleotide sequences, with mutation rates being observed for di- and tetra-nucleotide repeats at 0.001 mutations per generation (Goldstein & Pollock, 1997), which is up to four orders of magnitude higher than the mutation rate at non-SSR loci (Lacy, 1987).

High mutation rates mean that SSR loci may be polymorphic even in species otherwise characterised by low levels of genetic diversity. In highly inbred soybean (*Glycine max*)

cultivars; for example, Rongwen *et al.* (1995) reported 11 to 26 alleles per locus and an average heterozygosity of 0.87 at 7 SSR loci, substantially exceeding that obtained with allozyme and RFLP markers. By virtue of their extreme polymorphism, SSR loci are widely considered ideal markers for forensic identification, paternity analysis, gene mapping, plant breeding, conservation biology, and population genetics (Weber, 1990; Rafalski & Tingey, 1993; Gupta *et al.*, 1996; Jarne & Lagoda, 1996). However, as for STS-PCR in general, the requirement of DNA sequence knowledge means that a substantial investment of time and money is required to develop these markers. For this reason, the development and application of SSRs in plants has mostly been restricted to a few of our most agriculturally important crops.

One way it might be possible to facilitate more widespread use of SSRs in plants is to transfer SSR loci among congeners and across genera. In general, there appears to be moderate to complete cross-species transferability within genera (50 to 100% success), but successful cross-generic amplification appears to be much lower (see review in Peakall *et al.*, 1998). Nevertheless, the increasing numbers of SSR primer sequences published, may provide an additional source of markers, particularly among closely related taxa.

		Isozyme/ Allozyme	SSRs	Intron-PCR	PCR-SSR	Probed RFLP	PCR-product digest	RAPDs	AFLPs	Fragment sequence
Marker characteristics	Variability[a]	+	+++	++	++	+	+	++	+++	+
	Reproducibility[b]	++	+++	+++	+++	++	+++	++	+++	+++
	Technical ease[c]	+++	+	+	++	+	+	+++	++	++
	Precision[d]	++	++	+++	++	++	++	+	++	+++
	Development time[e]	Short	Long	Long	Medium	Medium	Medium	Short	Short	Medium
	Lab. equipment costs[f]	Low	Med.-High	Med.-High	Med.-High	High	Med.	Low	High	High
	Ongoing assay costs[g]	Low	Low-Med.	Med.	Med.-High	Med.	Low-Med.	Low-Med.	Med.-High	Med.-High

Table 2.1 Summary of the qualitative characteristics of various DNA markers (adapted from Rafalski and Tingey 1993, Karp and Edwards 1997, Powell *et al.* 1998 and Parker *et al.* 1998).

[a] Inherent capacity of a marker to reveal variation (see Powell *et al.* 1998). [b] Reproducibility refers to the ability to obtain the same genetic result for the same sample in repeated assays.[c] The level of skill required to obtain accurate genetic data once assay has been developed. [d] Differs from 'Reproducibility' in that a reproducible genetic result may not accurately reflect the diversity present within the sample due to effects such as epistatic effects with RAPDs (Heun & Helentjaris, 1993) or failure to amplify alleles (null alleles in SSRs). It can also be difficult to assess this category because different loci or alleles obtained from the same technique can also vary. [e] The time required to develop genetic assays and procedures for optimal performance. This can be shortened if primers/probes can be transferred from close relatives. [f] The financial requirements to prepare a laboratory to a satisfactory level for such analyses. These can differ depending on the separation and visualization processes chosen for the particular marker. [g] The financial requirement to obtain results after laboratory establishment. This too can vary depending on the specific techniques employed (radioactive vs. fluorescent labeling, agarose v acrylamide elctrophoresis).

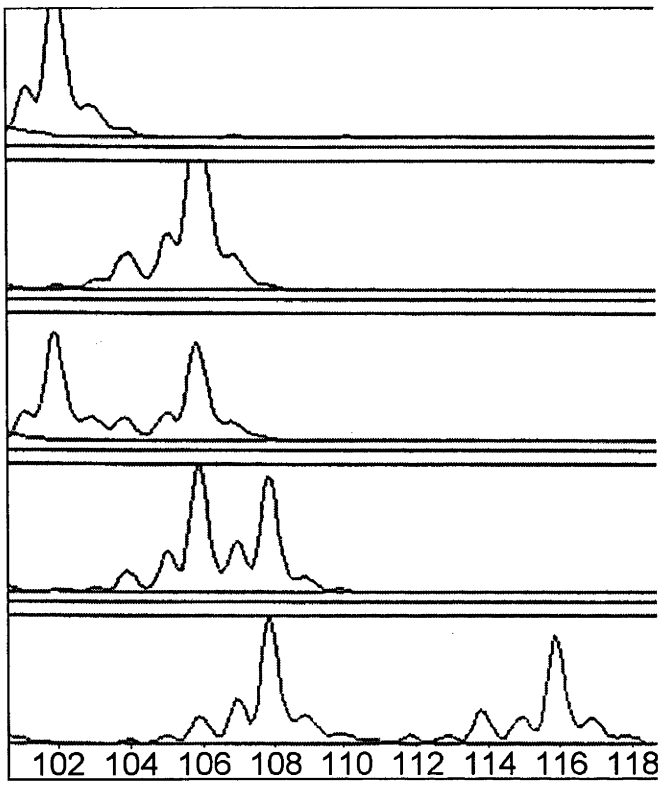


Figure 2.1 An example of a high resolution codominant DNA profile at a plant SSR locus as visualised on an automated DNA sequencer. Four alleles of a single locus are shown at sizes 102bp, 106 bp, 108bp and 116bp. A homozygote is revealed as a single peak and a heterozygote produces two peaks.

One technical complication for SSRs is that high resolution electrophoresis is required, particularly for di-nucleotide repeat SSRs where alleles may differ by only 2 base pairs. It is theoretically possible to achieve this level of resolution with high resolution agarose but many SSR loci may require sequencing-quality electrophoresis (Fig. 2.1). Some SSR loci/alleles also show characteristic sub-banding (stutter bands), which can complicate scoring. Furthermore, size homoplasy (false equality of alleles based on independent mutation to the same size) due to complex mutational processes within and beyond SSR regions can complicate the interpretation of SSR length data (Doyle *et al.*, 1998; Peakall *et al.* 1998). In particular, genetic relationships among taxa are likely to be seriously underestimated, a problem that is likely to be accentuated with increasing divergence (Peakall *et al.*, 1998). For studies concerned with defining genetic relationships among species, it may be impossible to infer SSR number from allele size.

Consequently, DNA sequencing of SSR alleles will be essential to minimise the risk of misinterpretation and to maximise the genetic information that can be obtained (Peakall *et al.*, 1998).

AFLPs

Amplified Fragment Length Polymorphism (AFLP) is another approach to PCR that combines the features of both STS and arbitrary-primed PCR. AFLPs are based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA (Vos *et al.*, 1995; Lin & Kuo, 1995). The technique involves 3 major steps: (1) restriction digestion of the genomic DNA, (2) ligation of short pieces of DNA (oligonucleotide adapters) to the ends of DNA fragments, and (3) high stringency selective amplification of a subset of the 'adapted' fragments, using primers of complementary sequence to the ligated adapter, plus 1-4 additional arbitrary nucleotides. The use of selective nucleotides is necessary to reduce the complexity of the profile by 'fishing out' only a proportion of the possible fragments. For example, the use of a single selective nucleotide on one of the primers will (on average) reduce the number of possible products by 1/4. As AFLPs are a subset of the entire genome digest, they are distributed throughout the genome (Maheswaran *et al.*, 1997). Compared with arbitrary-primed PCR such as RAPDs, AFLPs are performed under high stringency, and are therefore theoretically less sensitive to reaction conditions but, like other arbitrary-primed PCR procedures, AFLP produces a multi-locus fingerprint with polymorphism being apparent as either band presence or absence (Fig. 2.2).

Despite being a new technique, AFLPs have been widely adopted (e.g. Powell *et al.*, 1996; Sharma *et al.*, 1996; Roa *et al.*, 1997; Andrews *et al.*, 1998; Krauss, 1999). Such

popularity is largely due to the high variability of AFLPs. Studies in soybean (Powell *et al.*, 1996) and lentils (Sharma *et al.*, 1996) have found that AFLPs reveal more polymorphisms per primer than RAPDs. Furthermore, in a comparison of AFLPs, RFLPs, SSRs and RAPDs in barley, AFLPs revealed the greatest amount of genetic diversity (Russell *et al.*, 1997). Even in genetically depauperate natural populations of *Persoonia mollis*, AFLPs have been used to unambiguously assign paternity for 100% of the progeny collected (Krauss, 1999).

Unlike RAPDs, AFLPs are amplified with a relatively high temperature primer-annealing stage, resulting in low mismatch annealing and stringent PCR (Vos *et al.*, 1995). Consequently, AFLPs may avoid the problem of poor reproducibility from run to run and laboratory to laboratory sometimes reported for arbitrary primed methods (Heun & Helentjaris, 1993; Penner *et al.*, 1993, Vos *et al.*, 1995). In a comparison of the reproducibility of SSRs, RAPDs and AFLPs between different laboratories using the standardised PCR conditions, SSRs were reproducible though band sizes were slightly altered between laboratories (Jones *et al.*, 1997). For the multi-locus techniques, AFLPs produced a total of 172 fragments, of which only one was inconsistently amplified in one laboratory out of 6. On further analysis, it was found that DNA template variation was responsible. This result was far more reproducible than that obtained for RAPDs. From a total of 42 RAPD fragments, only 10 were amplified across all nine laboratories. While many bands were absent from only a single laboratory, most of this variation was due to inconsistent amplification across several laboratories. Such results show that the highly variable and stringent AFLP method is an attractive alternative. However, AFLPs have several limitations.

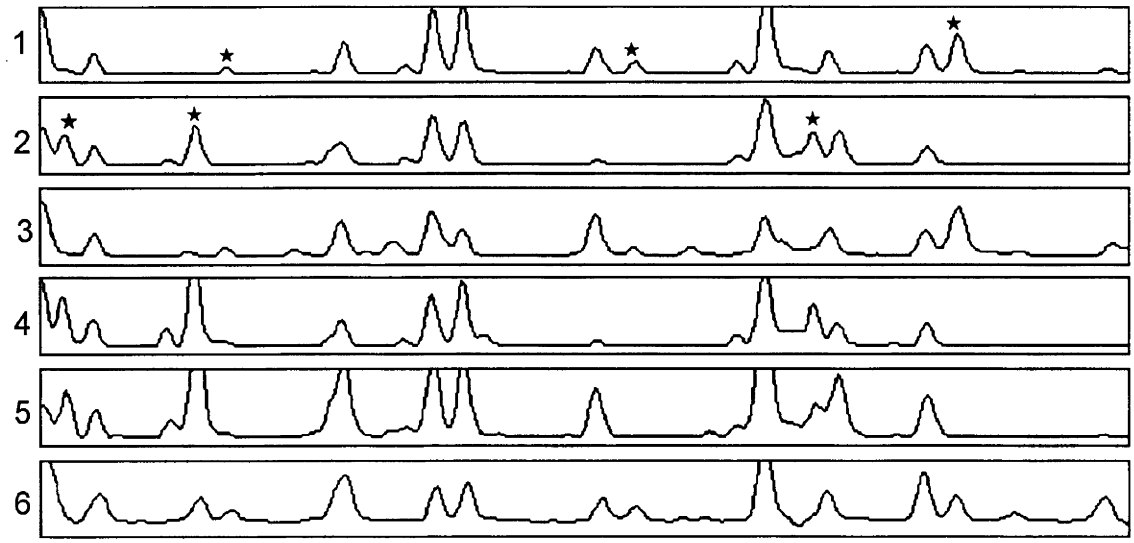


Figure 2.2 An example of dominant multilocus AFLP profiles in *Onopordum* thistles as visualised on an automated DNA sequencer. Panes 1-2, differences between *O. acanthium* and *O. illyricum* sampled from Europe. Panes 3-4, differences between *O. acanthium* and *O. illyricum* sampled from Australia, and panes 5-6, hybrid thistles sampled from Australia. Stars denote fragments diagnostic in parental species that segregate in hybrids. With this method, it is not possible to distinguish which peaks represent alleles or loci.

One technical consideration with AFLPs is that high resolution electrophoresis and radioactive, chemiluminescent or fluorescent labels are required to visualise the fragments (Fig. 2.2), which typically range in size from 50 to 500 bases (Lin & Kuo, 1995). Another consideration is that incomplete digestion of the DNA template in the first step of the AFLP procedure can lead to false polymorphism (Lin and Kuo, 1995). Thus, it is essential to ensure complete digestion, which may require the use of higher quality DNA than necessary for PCR in general. Finally, as with other multi-locus procedures, the resultant fingerprint can be complex, and it may be difficult to interpret band presence or absence for fragments which are sub-optimally amplified in individual reactions. It is also not possible with AFLPs to discriminate between alleles and loci based on size. Consequently, it is not possible to differentiate between heterozygotes and homozygotes for individual loci without conducting inheritance studies. In this sense, AFLPs provide less genetic information per locus than codominant markers such

as SSRs, but this is somewhat offset by the large number of loci that can be readily screened.

Other methods

PCR SSR identification

Genetic marker refinement and development is ongoing. Presently a series of methods that attempt to visualise SSR variation without the time consuming development costs are emerging. These include 'Inter-SSR-Amplification' (ISA) (Zietkiewicz *et al.*, 1994), 'Selectively Amplified Microsatellite Locus' (SAMPL) (Witsenboer *et al.*, 1997), 'Microsatellite-Primed PCR' (MP-PCR) (Sharma *et al.*, 1995) and the procedure for producing 'Random Amplified Microsatellite Polymorphisms' (RAMPs) (Wu *et al.*, 1994). All these methods produce multi-locus profiles and their advantages and disadvantages over other methods remain to be determined.

Intron PCR

Intron-PCR is the newest class of STS marker to emerge. Like SSRs, intron-PCR reveals polymorphic codominant markers. As introns are non-coding, they are largely selectively neutral (though factors such as length and secondary structure may be subject to selection). As a consequence, they have relatively high mutation rates (Friesen *et al.*, 1997) yielding high levels of polymorphism. For example, in comparison with RFLP, and SSR markers linked to SAD genes in sunflower, introns were found to be nearly twice as polymorphic as RFLPs and slightly more polymorphic than SSRs (Hongtrakul *et al.*, 1998).

An additional benefit to the high variability of introns may be that as the flanking regions are conserved, cross transferability of intron-PCR primers may be more widespread in plants than SSR primers (Hongtrakul *et al.*, 1998). As primers are designed to amplify within coding regions which are conserved through selection, the priming sites are likely to be less variable than for SSRs (SSR flanking regions can also be highly variable, see Peakall *et al.*, 1998). Mutation at intron sites may be less complicated than for SSRs in which length polymorphism occurs through slippage of repeat motifs (Goldstein & Pollock, 1997). As a consequence, introns may be more useful for reconstructing relationships between taxa than SSRs since the interpretation of intron variation may not require complex assumptions about the underlying mutational processes.

Introns can vary both in length and sequence. Several electrophoretic techniques are available for intron-PCR, though most are technically demanding. While length variation may be revealed by traditional electrophoretic techniques, sequence variation which may account for much of intron polymorphism, can only be revealed by specialised techniques. Currently, single strand conformational polymorphism electrophoresis (SSCP) (Friesen *et al.*, 1997), or Temperature Gradient Gel Electrophoresis (TGGE) (Heslewood *et al.*, 1998) are the two main approaches. SSCP electrophoresis relies on the differential folding patterns of single stranded DNA, resulting in mobility differences whereas TGGE relies on the differential denaturation of double stranded DNA based on small sequence differences within the gel.

Sequencing

Ultimately, all genetic markers seek to identify sequence variation. In this regard, sequencing DNA is the ultimate marker and for this reason has been popular for many years. Furthermore, sequencing reactions are highly reproducible, and results are easy to interpret (Fig. 2.3). However, in the past, sequencing was limited to well characterised genes and spacer regions which were generally not variable at the species level. Nonetheless, many more highly variable regions are now becoming well known, each of which has universal primers, allowing amplification in most plant species. Universal primers for mtDNA and cpDNA have been published (Dumolinlapegue *et al.*, 1997) and the number of known regions is rapidly growing. In addition, markers derived from other techniques can now be sequenced more simply and cost-effectively than in the past. Consequently, where variable regions can be identified, sequencing these products may provide an additional tool for weed research.

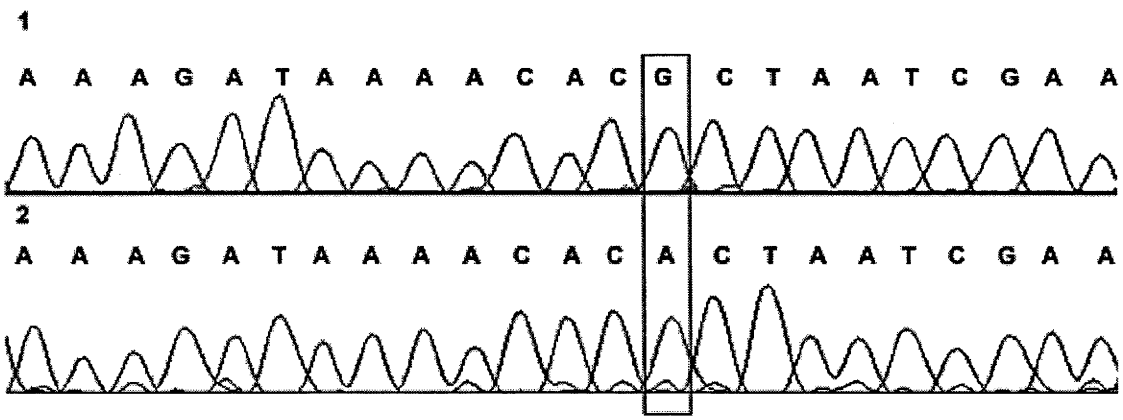


Figure 2.3 An example of nucleotide sequence variation in two thistles for the trnL-trnF inter-genic spacer visualised on an automated DNA sequencer. Variation is revealed by different nucleotide identities at the same position. Variation can also be represented by 'indels' (insertions and deletions) after DNA sequence alignment.

Applications of PCR based genetic markers in weed ecology

Genetic study is becoming increasingly widespread in weed research. Since the review of Nissen *et al.* (1995), there has been a growing number of studies applying RAPDs

and RFLPs, but as yet, the application of SSRs and AFLPs has been limited to only one paper for SSRs (Mörchen *et al.*, 1996) and two for AFLPs (Andrews *et al.*, 1998; O'Hanlon *et al.*, 1999). However, regardless of the marker used, the goals of genetic weed research can be broadly classified into three areas of investigation: (1) patterns of genetic diversity within invading weeds, (2) the taxonomic identity of weeds and, (3) determining the origin/s of introduced weeds. By way of illustration, here I briefly describe studies employing genetic markers since the review of Nissen *et al.*, (1995).

Patterns of genetic diversity in invasive populations

There is a widespread concern that weed species with higher levels of genetic diversity will exhibit considerable potential for weed adaptation, and therefore may be able to reduce the effectiveness of weed control (Dekker, 1997; Holt & Hochberg, 1997). As such determining the magnitude of variation in weedy species is often considered a high priority. Despite these efforts, there is a lack of evidence correlating genetic diversity with physiological, morphological or other ecological characteristics such as susceptibility to biocontrol agents (Chaboudez & Sheppard, 1992). Consequently, this is an area of important and ongoing research (Dekker, 1997; Nissen *et al.*, 1995).

RAPDs were employed by Moodie *et al.*, (1997) to investigate patterns of temporal and population genetic variation in *Sinapis arvensis* L., and how these patterns were influenced by herbicidal control. Unfortunately sampling within populations was minimal, and rigorous comparisons of the magnitude of within population components of variation were not possible. Nevertheless, preliminary data suggested that populations with a history of herbicidal control showed no less variation than untreated populations. Preliminary data also suggested that in wild mustard populations, patterns

of genetic diversity may change annually, and that to estimate genetic diversity accurately, sampling over multiple years may be required. Even though it was difficult to determine within population variance, most individuals from different populations were extremely divergent from one another with as low as 15% genetic similarity being observed. The authors concluded that RAPDs were an extremely useful marker system for use in wild mustard populations. In another assessment of new marker technology, Morchen *et al.*, (1996) used SSRs in cultivated, wild and weedy beets (*Beta vulgaris* L.). SSR variation was shown to differ significantly between these groups. Wild beets were the most diverse, cultivated beets were the least diverse and weedy beets were intermediate between these two groups. The three SSR loci identified were highly variable, containing 11.6 alleles on average and the patterns of relatedness between genotypes was similar to that obtained with RFLPs. However, as SSRs were technically less demanding, more cost effective and easier to score than RFLPs, they were identified as the marker of choice for future genetic studies in beets. By having access to such a powerful molecular marker system in taxa with wild, weedy and cultivated representatives, SSRs may help inform previous studies of the evolution of weedy beets from cultivated and wild representatives, and genetic introgression into weeds from transgenes (Boudry *et al.*, 1993).

Apart from broad descriptions of diversity, molecular tools can be used to target patterns of diversity in relation to specific phenotypic traits. In the first published use of AFLP markers in weed research, Andrews *et al.* (1998) identified a correlation between the genetic identity of *Avena fatua* L. individuals, the field from which the individuals were sampled, and ACCase inhibitor resistance. Such studies employing combinations of highly polymorphic markers, ecological data and spatial analyses are extremely

promising for identifying the genetic basis of specific phenotypic traits in populations and the spread of undesirable traits.

Breeding systems

Breeding system, often inferred from patterns of population genetic structure, has been shown to be an important predictor of biocontrol success, as recombination allows for the development of genotypes that are difficult to control (Dekker, 1997). Various genetic markers have proved useful for investigating breeding systems through progeny analysis or population genetic structure (Nissen, 1995). Traditional genetic markers such as isozymes remain the most popular technique for achieving this goal, with twelve of the seventeen published papers in weed genetics between 1995 and 1998 relying on them (e.g. Michalakis *et al.*, 1993; Wang *et al.*, 1995; Morrison & Scott, 1996; Runyeon & Prentice, 1996; Strefeler *et al.*, 1996; Sun, 1997; Meerts *et al.*, 1998). However, RAPD markers have also been used by Okoli *et al.*, (1997) for investigating population structure in *Cyperus rotundus* L. and *Cyperus esculentus* L. The study found that purple nutsedge was much less diverse than its congener with the difference between species most probably reflecting contrasting breeding systems. Species that display high divergence between populations but low within population variation are characteristically inbreeders, with the converse being characteristic of outbreeding species. However, these patterns are prone to the confounding effects of historical processes. Wild mustard is a self incompatible outcrosser that displays low levels of within-population genetic diversity but high differentiation between populations (Moodie *et al.*, 1997), perhaps due to the widespread sampling regime employed. Progeny analysis is more powerful in determining the breeding system of weed species.

The study of breeding systems is also crucial to understanding the likelihood and consequences of crop-weed gene flow (Darmency, 1994). Boudry *et al.* (1993) have shown that gene exchange between cultivated and wild beets is common. Furthermore, many weedy beets appear to be derivative of cultivated beets, through accidental pollination. Given that herbicide-resistant transgenic beets are capable of interbreeding with weedy types, herbicide resistance may unintentionally be introduced into weedy populations. Another example of evidence for gene flow between crops and weedy relatives can be found in *Sorghum halepense* L. and crop sorghum (*Sorghum bicolor* (L.) Moench). A single variable isoenzyme locus revealed widespread introgression of crop sorghum genes into johnsongrass (Arriola & Ellstrand, 1996). Such studies on the capacity for gene flow between close relatives has provided critical data on the biosafety of released transgenes. However, the ongoing consequences of such hybridisation need be quantified. Where transgenes differ only slightly from wild relatives, and where particular genic traits confer fitness advantages to weeds, hypervariable and STS-PCR may provide additional important evidence.

Taxon identification and phylogenetic relationships

Recently, O'Hanlon *et al.*, (1999; see also *Chapter Five* and *Chapter Six*) revealed widespread hybridisation and introgression in invading *Onopordum* in south-east Australia. Using only a single AFLP primer pair combination, a total of 143 AFLP fragments were identified, 108 of which were polymorphic. Fragments diagnostic for parental species segregated in many individuals confirming that, in Australia, *Onopordum* exists as a species complex capable of inter-species genetic transfer. Such a finding has potentially important implications for the control of *Onopordum*. Given that

the parental species in this complex exhibit ecological differences such as perennation in *O. illyricum* L., a smaller head size and more pubescent leaves and stems in *O. acanthium* L., the potential for gene exchange between these species may increase the potential for adaptation in the group in response to particular herbicidal, cultural and biocontrol strategies (O'Hanlon *et al.* 2000)[b]. Clearly, further research is required to explore these possibilities.

Using seven RAPD primers, Abad *et al.*, (1998) were able to discriminate between ten individual clones of *Cyperus esculentus* L. including several weedy and cultivated clones. Weedy clones were all closely related to, and distinct from all but two cultivated clones. It is of interest that "weediness" appears to be distributed phylogenetically in this group. If "weediness" has a genetic basis, such research may help provide a genetic understanding of what, if anything, differentiates plants with a high potential of becoming weeds from those with a low one. Furthermore, this analysis warns of the weedy potential of these two cultivated clones in an alien range. RAPDs appear to have been useful in reconstructing evolutionary relationships in this example, however, several criteria must be satisfied in order to confidently reconstruct phylogenetic relationships with accuracy (see Swofford *et al.*, 1996).

One feature common to all genetic analyses markers that rely on scoring differences based on gel separation is the assumption of the equality (homology) of co-migrating fragments. Katzir *et al.*, (1996), when performing phylogenetic analysis on *Orobanche* species, confirmed that co-migrating RAPD fragments were homologous using southern-hybridisation analysis. Such confirmation of homology is particularly important when using RAPDs (Rieseberg, 1996) as size homoplasy can be common. AFLPs and SSRs, as highly polymorphic markers, may also require such confirmation

of homology between co-migrating fragments for diverged taxa. While for SSRs this level of divergence may equate with boundaries between genera (Peakall *et al.*, 1998), it is unclear how AFLPs behave across taxonomic boundaries.

Geographic origin/s

Another finer scale taxonomic endeavour in weed research is in determining the geographic origin/s of invading weeds. Where differentiation between populations has occurred in the native range, it is theoretically possible to pinpoint the source of introduction of weedy material. However, most studies to date have focussed on broad descriptions of the source and describe only the continent or country of origin. As the level of genetic differentiation within species distributed across geographic regions is usually lower than may be found between species, the discriminating power of selected markers becomes an important issue. Differentiation between taxa may be revealed by allele frequency differences or fixed diagnostic differences. However, the use of allele frequency differences may be problematic. The relatively narrow genetic subset of the taxon's entire genetic diversity that is usually introduced automatically confounds comparisons of allele frequencies between the native and alien range. When a few individuals from a single population are introduced, and a bottleneck is imposed, it is unlikely that allele frequencies will remain at their original levels. In contrast, the use of diagnostic genetic differences does not suffer from this difficulty. Provided that such differences can be identified, they are more powerful in the reconstruction of patterns of invasion. However, the identification of sufficient numbers of diagnostic markers requires greater degrees of variation to be revealed, either through using more highly

variable markers, or through screening of more loci. In addition, to pinpoint geographic origins, sufficient sampling of individuals from all parts of the native range is required.

Rowe *et al.*, (1997) applied a combination of RAPDs and chloroplast RFLPs to determine that *Euphorbia esula* L. in North America comprise populations with surprisingly high levels of cpDNA variation. Such high variation was interpreted as either high inherent variability in the species native range, and/or multiple introductions into the alien range. However, given the limited sampling of Eurasian populations, it was impossible to determine which of these possibilities was most likely. As this question may inform decisions relating to selecting biocontrol agents (Rowe *et al.*, 1997) this study serves as a basis for future research. It also further illustrates the complexity of genetic marker interpretation because, not only was leafy spurge highly diverse, but the diversity was spatially distributed. The degree of relatedness was correlated with the spatial proximity of the samples assessed. Furthermore, the two types of molecular markers employed, RFLPs and RAPDs, showed different patterns of spatial structure. The RFLP data were derived from the maternally inherited chloroplast, whereas RAPD fragments were generated from the entire genome, which is mostly inherited biparentally. Therefore, as pollen which usually contains no chloroplast, and seed have a different mechanisms and rates of dispersal, RFLP and RAPD data may produce different geographic structures.

Using sequence data from the nuclear region ITS1, Scott *et al.* (1998) showed that the likely source of introduction of *Chromolaena odorata* (L.) King & Robertson in Australia was Brazil. A population of this weed that flowered at an earlier time than the other populations in Queensland was found to differ by 4 base pairs from later flowering populations. This genotype was found to occur elsewhere only in Brazil, where the

other genotype was also present. Even though ITS1 is a highly variable region, DNA sequence of regions such as this usually reveal only low levels of variation. Therefore, it was somewhat fortuitous that this diagnostic difference was identified, as the amount of sequence variation was far lower than the phenotypic variation displayed by the species (Scott *et al.*, 1998). In general, the sequencing of hypervariable markers such as introns or SSRs may be more helpful for studies of weed phylogeography.

Note on statistical analysis

(superscripts refer to Fig. 2.4)
Several statistical procedures[1] for analysing multi-allelic codominant data are well established and have been reviewed on several occasions (Berg & Hamrick, 1997; Brown & Weir, 1983). Recently, a suite of procedures has been developed to estimate genetic parameters for the same allelic data, many of which utilise permutational or Markov-chain procedures to test significance[1] (Rousset & Raymond, 1997). In parallel, a framework for estimating population genetic parameters using allelic data as well as binary multilocus band data and DNA haplotypes have been developed. At its core is the pairwise genetic distance matrix[3]. By estimating between-sample distances, analyses of population structure and differentiation via Analysis of Molecular Variance (AMOVA) as well as spatial autocorrelation and related techniques[5] can be performed (Excoffier *et al.*, 1992; Peakall *et al.*, 1995; Smouse & Peakall, 1999). The matrix also forms the basis for performing multivariate analyses[7] (O'Hanlon *et al.*, 1999) and phylogenetic clustering procedures[8] (Abad *et al.*, 1998).

Numerous techniques for phylogenetic reconstruction[8] have been developed (Swofford *et al.*, 1996). Traditional methods such as distance/similarity hierarchical tree building and cladistic techniques may be applied to most genetic data. However, multilocus band

data can still be prone to statistical non-independence (Rieseberg, 1996), and hence can be problematic for phylogenetic analyses. In contrast, probably the most suitable genetic marker for phylogenetic analysis is DNA sequence data. Consequently, it has long been used for this purpose, and newer more sophisticated techniques relying on a knowledge of the mutational behaviour of nucleotides, such as maximum likelihood, have been developed (Swofford *et al.*, 1996).

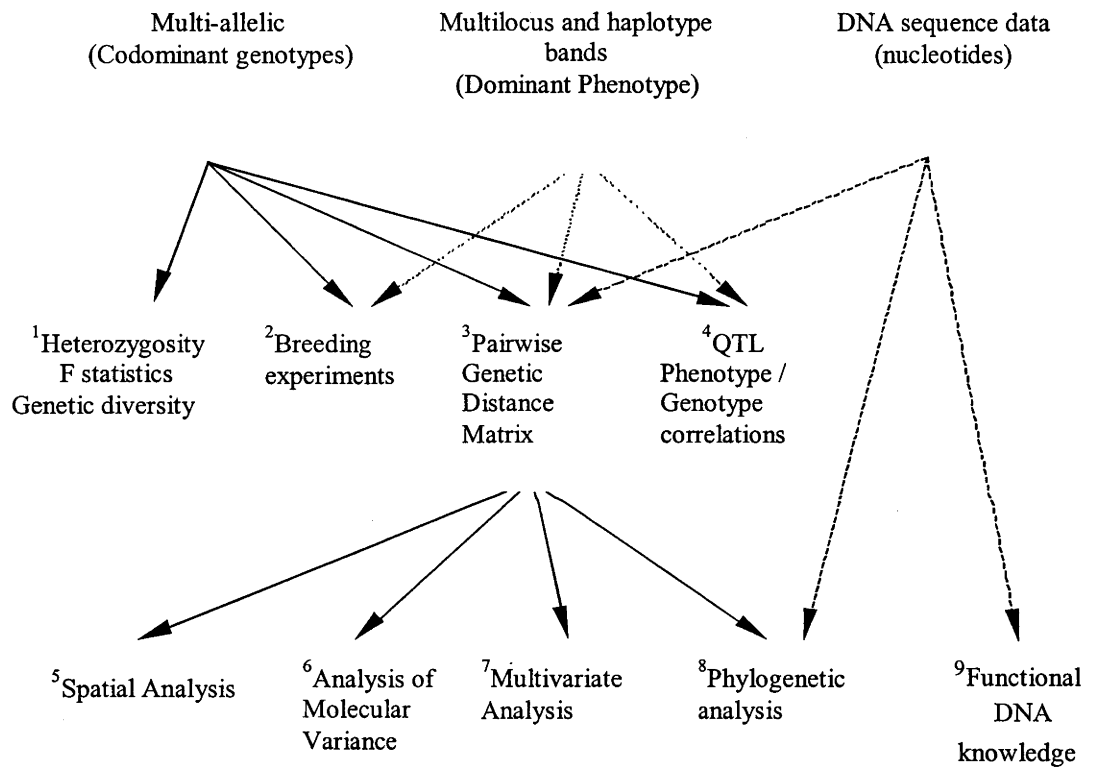


Figure 2.4 A simple framework showing the statistical options available for use with genetic data in weed ecology. Superscripts refer to text in section "Statistical developments".

Any of the three data types can be used in performing studies relating phenotype to genotype[4] (Andrews *et al.*, 1998) by means of phenotype linked markers (QTLs - quantitative trait loci) or in studying breeding systems[2] through direct analysis of progeny (Brown, 1990; Milligan & McCurry, 1993). However QTL analysis usually requires many markers distributed throughout the genome (Kearsey, 1998) which

cannot be produced through sequencing, and breeding system analysis typically requires more variation than revealed by sequencing. Nonetheless, nucleotide sequence itself can reveal the function of ecologically significant genetic markers[9] and as databases with such information grow in sophistication, such utilities may become more widely adopted in weed ecology.

Conclusions

Marker technology

Recent developments in molecular techniques give the weed scientist a larger array of genetic tools than previously. These new markers show a greater capacity to uncover genetic variation and hence give greater resolution in questions of interest. Moreover, these techniques, once established, may be cost effective, less technically demanding, and more reliable than their predecessors. Such improvements may allow for more accurate resolution of species complexes and breeding systems, identification of the origins of weed material, particularly where multiple introductions have occurred, as well as defining genetic relationships between closely related taxa. Despite these benefits, existing techniques will continue to serve as effective markers for many questions. Methods such as isozymes and RAPDs in particular, are still very popular due to their low cost, and rapid turnaround. These older techniques may also provide valuable preliminary data for pilot studies and exploring initial hypotheses. The newer and often more expensive techniques may therefore be most appropriately applied to species of low genetic diversity, when sufficient resolution is not provided by other markers, or where the studied taxa have close relatives that have had STS-PCR techniques optimised for use with them.

Applications to weed research

Descriptive studies of patterns of genetic diversity in weedy populations rarely alter management practices. However, they can be extremely important as they often provide essential background for further, more focussed research. Unfortunately, this kind of research does not always eventuate. In contrast, genetical studies that are designed to investigate specific phenotypes such as herbicide resistance and host specificity, or other ecological processes such as gene flow generally have more clearly definable outcomes and hence often provide more useful information.

Phylogeny and taxonomy are well established fields that can generally be conducted without the same intensive sampling that population analyses sometimes require. Where there is taxonomic ambiguity, the accurate identification of weed taxa can assist in their control through the accurate dissemination of ecological information specific for those taxa. Moreover, the understanding of evolutionary relationships between species and sub-specific taxa can inform choices regarding biocontrol agent choice, as well as in helping obtain information from close relatives where it cannot be obtained for the weed species itself. In addition, phylogeny may assist in identifying plants with a propensity toward "weediness" in certain habitats.

While it is often difficult to pinpoint the source of invading weeds, when successfully achieved, there may be considerable benefits. Apart from the benefits of having an accurate taxonomic identification and an understanding of evolutionary relationships, locating sources of origins may enable a better understanding of how colonisation occurred or continues to occur. However, due to the highly variable markers required and the large number of samples that may be required, studies of this kind may require a

heavy investment of resources. Plant taxa that are clonal or of low intrinsic diversity may be most amenable to studies of this kind.

Implications for studies in Onopordum

Before beginning genetic research within the *Onopordum* thistles, several lines of inquiry were apparent. Firstly was the taxonomic evaluation of *Onopordum* thistles in Australia. Such studies would require extensive sampling within the Australian range and representative collections of various species from the native range. Such a collection would potentially enable the identification of species diagnostic markers with which to identify Australian thistles. For this purpose, anonymously primed techniques particularly AFLPs, appear to be an extremely promising avenue. SSRs, while also suitable, have not been developed within close relatives of *Onopordum* and consequently may require substantially more time and expense than AFLPs. Additionally, provided enough AFLP variation is identified, this may enable the geographic origins of introduced material to be identified. While AFLPs may also provide some information on relationships, other markers such as sequencing of chloroplast regions may provide more phylogenetically informative data on which to base analyses. However, the two types of data may also be used in concert, thus providing additional resolution as well as the capacity to differentiate between maternal and paternal contributions, and thus disentangle potential reticulations. These questions form the basis of the thesis that follows.

Chapter Three *A simple method for the detection of size homoplasy in PCR based multilocus fingerprints* [2]

The co-migration of non-homologous fragments (size-homoplasy) is a common feature of anonymously primed multi-locus fingerprinting techniques. Here I demonstrate that a modification of the AFLP procedure can be used to identify incidences of size homoplasy and calculate the frequency of size homoplasy across a range of taxonomic boundaries within the Carduinae thistles.

PCR based methods that produce multilocus DNA profiles such as randomly amplified polymorphic DNA (RAPDs) (Welsh & McClelland 1990, Williams *et al.* 1990) and amplified fragment length polymorphism (AFLPs) (Lin and Kuo 1995, Vos *et al.* 1995) have become widely adopted tools for systematic and ecological applications (Hadrys *et al.* 1992, Mueller and Ullrich, 1999). With these multilocus techniques, it is not possible to directly distinguish between loci and alleles in DNA profiles. Consequently, comparative studies rely on the assumption that co-migrating fragments are homologous, and that when two samples share such a fragment, they share a genetic similarity.

However, this assumption is not always valid. Size homoplasy (the co-migration of fragments of the same size but different identity) can result in a false interpretations of genetic similarity (Dowling *et al.* 1996, Rieseberg 1996, Peakall *et al.* 1998). The risk of such size homoplasy is particularly high in multilocus DNA profiles. In an analysis of 220 co-migrating RAPD fragments in a wild sunflower species complex, a combination of southern hybridization and RAPD fragment digestion revealed that only 79.1% were homologous (Rieseberg, 1996). A combination of size homoplasy, gene

[2] This chapter is an expansion of O'Hanlon & Peakall (2000)

duplication, and codominant inheritance were considered to be factors that contributed to the frequency of non-homology. Furthermore, in a comparison of 47 fragments of similar but different size, 42 could be traced to the same genomic location, indicating that they were different alleles of the same locus. Thormann *et al.* (1994) showed that all of the 15 fragments that they tested for homology by southern hybridization within 6 *Brassica* and one *Raphanus* species were homologous. In contrast, when comparing between species, 3 of these 15 fragments were not homologous.

Southern hybridization and fragment digestion can be used to identify both size homoplasy and the differential migration of alleles at the same locus. However, for many molecular ecology laboratories, this is technically demanding and time consuming. As a consequence, studies that employ multilocus methods rarely investigate size homoplasy and homology. In this technical note I introduce a simple PCR based approach for detecting size homoplasy. I demonstrate this approach using the AFLP method which is particularly well suited for my purpose. However, the approach can in theory be extended to any PCR based multilocus method including RAPDs.

In the AFLP method, two different primers each beginning with sequences complementary to ligated adapters, followed by (usually 3) arbitrary selective nucleotides, are used to amplify multilocus fragments. In general, an MseI primer with 3 selective nucleotides (MseI-3) and an EcoRI primer, also with 3 selective nucleotides (EcoRI-3) are used. The addition of extra nucleotides is essential as it allows the amplification of only a subset of the many fragments produced through digestion, and thus reduces the total number of fragments, and complexity in the AFLP profile. The

use of selective nucleotides offers the opportunity to examine size homoplasy in a novel way.

In addition to the primers with 3 selective nucleotides (MseI-CTA and EcoRI-ACA), I synthesized a set of MseI primers with one extra nucleotide (i.e. MseI-CTAA, MseI-CTAC, MseI-CTAG, MseI-CTAT). AFLP profiles produced from these 4 primers should be a subset of the profile obtained using the MseI-CTA primer, and by analyzing these profiles separately, the identification of the nucleotide adjacent to the 3' end of the MseI-CTA primer in these fragments may be possible. In this chapter I test this approach in taxa of different degrees of divergence. In addition, I calculate the frequency of shared fragments and size homoplasy across different levels of taxonomic divergence within the thistles.

Table 3.1 List of species studied within the Carduinae and their relevant Genbank accession numbers for the nucleotide sequence of the chloroplast trnL-trnF region.*Not submitted to Genbank but the same sequence for trnL-trnF.

Subtribe	Species	Genbank accession trnL-trnF
Echinopsidinae		
	Echinops sphaerocephalus L.	AF129830
Carduinae		
	Arctium lappa L.	AF129824
	Carduus nutans L.	AF129825
	Cirsium vulgare (Savi) Ten.	AF129826
	Cousinia hystrix C.A. Mey.	AF129827
	Cynara cardunculus L.	AF129828
	Cynara humilis L.	AF129829
	Galactites tomentosa Moench	AF129831
	Notobasis syriacus (L.) Cass.	AF129832
	Onopordum acanthium L.	AF129833
	Onopordum illyricum L.	AF129833*
	Picnomon acarna (L.) Cass.	AF129834
	Ptilostemon afer (Jacq.) Greuter	AF129835
	Silybum marianum (L.) Gaertn.	AF129836
	Tyrimnus leucographus (L.) Cass.	AF129837

I examined a range of thistles within the Carduinae (Bremer, 1994), thistles, with known and different levels of cpDNA divergence (Table 3.1) distance. I sampled broadly so as to obtain several comparisons: (1) between congeneric species with no cpDNA divergence, (2) between closely related genera with moderate cpDNA divergence and (3) beyond subtribes with high cpDNA divergence (Table 3.1). For MseI-CTA primed reactions, I performed AFLP analysis with standard reagents (see O'Hanlon *et al.* 1999 for details) and PCR conditions (1X (94°C, 60s: 65°C, 30s: 72°C, 60s), 10X (94°C, 30s: 65-1 °C until 56°C, 30s:72°C, 60s), 23X (94°C, 30s: 56°C, 30s: 72°C, 60s), 4°C). Vos *et al.* (1995) reported mismatch annealing when MseI-4 primers were employed with these PCR conditions. I therefore followed Krauss (1999) who used more stringent PCR conditions for MseI-4 primers with annealing temperature commencing at 70°C and decreasing by 1 °C each cycle until 61°C was reached, followed by 23 cycles. Products were electrophoresed in a 5% polyacrylamide denaturing gel for high resolution set up in an ABI 377 automatic sequencer, and resultant profiles were analyzed with ABI GeneScan 3.0.

The superimposition of the four MseI-CTAx profiles produced a nearly identical profile as that obtained with MseI-CTA (Figure 3.1). Of the 94 shared fragments produced by MseI-CTA, only 3 were amplified by more than one of the MseI-CTAx primers. Such a result confirms that mismatching occurred only rarely in the selective amplification with MseI-4 primers and the more stringent thermal regime employed with them (see Vos *et al.* 1995). I could therefore identify 91 of the nucleotides adjacent to the 3' end of the MseI-CTA primer and calculate the frequency of size homoplasy across taxa.

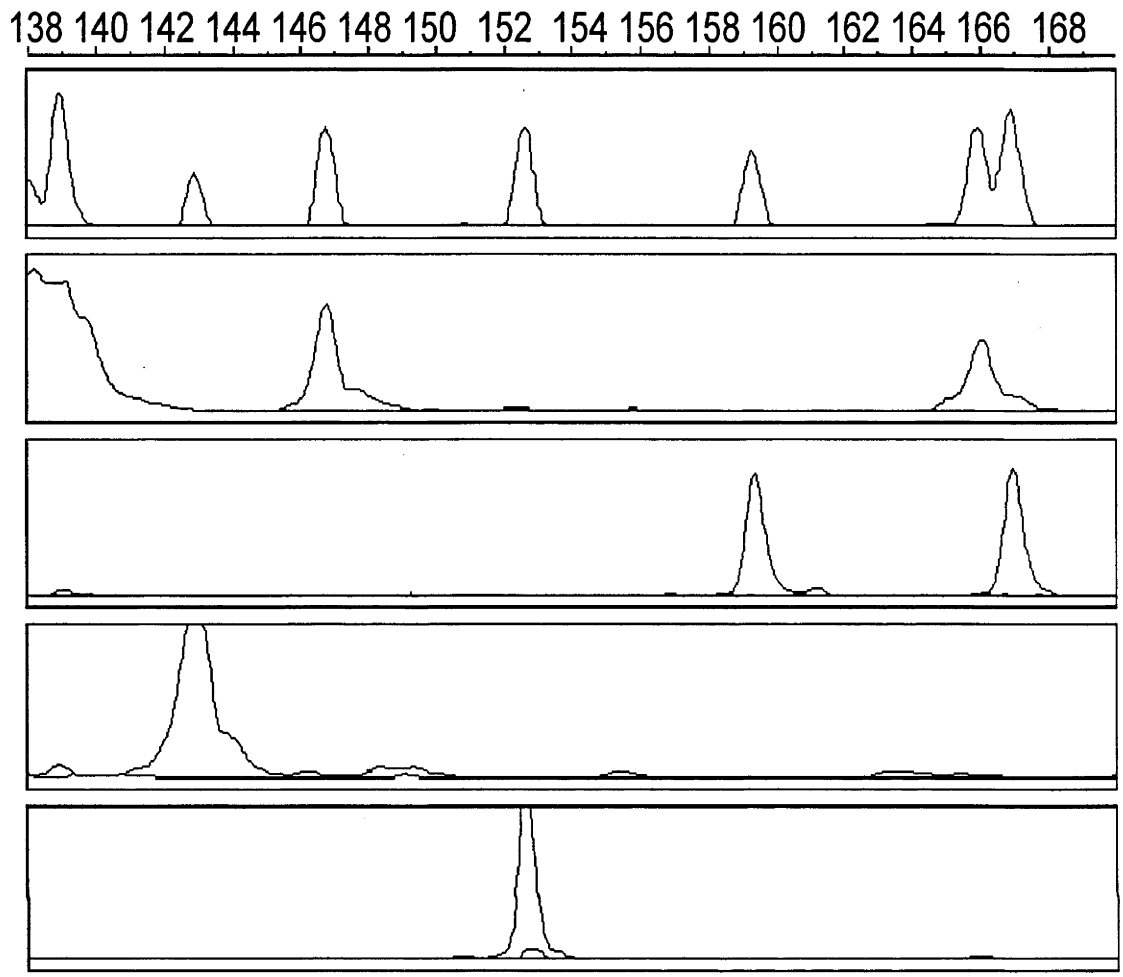


Figure 3.1 Comparison of AFLP profiles for Carduinae thistles derived from a primer with 3 selective nucleotides (Pane 1) and 4 primers with an additional selective nucleotide (Pane 2, A: Pane 3, C: Pane 4, G: Pane 5, T). The profiles produced from each of the extra selective primers are a subset of that obtained from the standard AFLP primer. The identity of the 4th nucleotide for each fragment in pane 1 is revealed in the remaining panes (i.e. 139bp, A; 143bp, G; 147bp, A; 153bp, T; 159bp, C; 166bp, A; 167bp, C)

Overall, of the 91 fragments shared between samples within this group, 53% of them had a different nucleotide at the analyzed position. For most cases this probably represents different loci co-migrating rather than the same locus having a different allelic states. While such a figure is extremely high, the frequency of size homoplasy was related to the taxonomic divergence between the samples (Figure 3.2). Between congeners, an average of 2.5% size homoplasy was observed, while between species from different subtribes a high of 88% was observed. Regression analysis indicated that

for these primer combinations, 100% of any shared fragments may be non-homologous in this group beyond cpDNA divergence of 0.3%.

In addition, there was a negative relationship between the number of fragments shared between samples, and the cpDNA divergence (Figure 3.2) though this relationship is probably biased as a consequence of the high number of shared fragments at low cpDNA divergence only. Increasing taxonomic divergence did not always result in fewer shared fragments, as predicted by Mueller and Ullrich (1999). For example, the number of shared fragments was sometimes relatively high between members from other subtribes (Table 3.2).

Table 3.2 Comparisons of the total number of shared AFLP fragments (lower triangular matrix), and the number of non-homologous but co-migrating fragments (upper triangular matrix). Lighter shaded cells represent comparisons across subtribes and darker shaded cells represent congeneric comparisons.

	Ech	*Jur*	*Cou*	*Arc*	*Not*	*Pic*	*Tyr*	*Gal*	*Sil*	*Car*	*Cir*	*Cy. hu*	*Cy. ca*	*O. ac*	*O. il*	*Pti*
Echinops	-	0	2	4	0	4	1	3	0	2	1	0	0	0	1	0
Jurinea	1	-	1	0	2	4	2	4	2	1	4	2	4	1	3	2
Cousinia	3	2	-	2	1	3	5	2	0	1	0	2	3	1	2	2
Arctium	7	4	3	-	2	2	2	0	3	2	6	2	4	1	2	5
Notobasis	0	2	2	4	-	1	0	2	4	1	6	3	4	0	2	4
Picnomon	4	4	3	5	6	-	3	4	4	2	5	3	4	3	4	2
Tyrimnus	1	6	7	3	0	3	-	1	1	2	1	2	0	1	2	0
Galactites	4	5	4	1	2	4	2	-	1	2	0	2	4	2	2	0
Silybum	0	3	0	3	5	9	3	4	-	2	2	1	2	1	0	1
Carduus	2	3	1	3	2	3	8	3	4	-	4	1	0	0	3	2
Cirsium	2	5	3	9	7	9	3	0	7	5	-	2	5	2	3	3
Cy. humilis	1	3	2	3	4	3	2	2	1	1	3	-	0	0	2	0
Cy. cardunculus	1	4	4	5	6	4	1	5	2	0	8	15	-	1	2	2
O. acanthium	1	3	3	4	1	5	4	2	1	1	4	1	3	-	1	0
O. illyricum	3	5	4	3	2	4	7	2	0	4	5	2	5	19	-	3
Ptilostemon	2	5	3	7	4	4	0	0	1	2	4	3	4	3	4	-

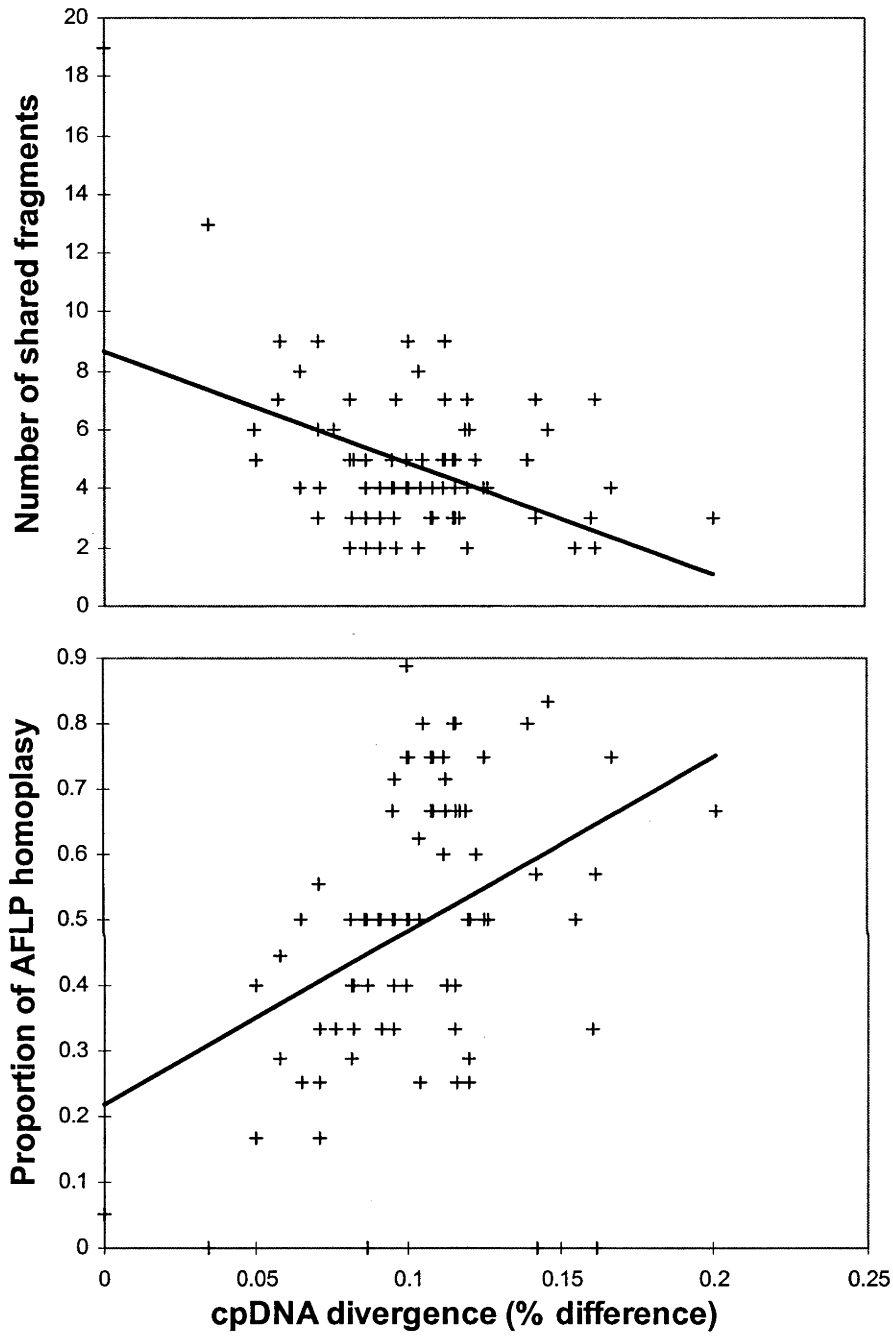


Figure 3.2 Relationships between (A) the number of co-migrating fragments and cpDNA divergence (Mantel correlation; -0.256, p=0.03) and (B) the proportion of shared fragments that are non-homologous and cpDNA divergence (Mantel correlation; 0.259, p<0.001) spanning a range of taxonomic divergence for the Carduinae thistles.

Furthermore, the difficulty in predicting the number of shared fragments is illustrated by the fact that the *Onopordum* species differed twofold in the number of shared fragments between the distantly related *Cousinia* and themselves. In addition, the two *Cynara*

species shared 13 fragments but the relatively distantly related *Cirsium* and *Cousinia* shared a relatively high 9 fragments. Therefore, size homoplasy was and will be a greater problem than the identification of sufficient numbers of shared fragments.

These results are conservative as, by chance, 25% of shared non-homologous fragments will have the same nucleotide at the fourth position. This percentage could be reduced by the use of primers with additional nucleotides (e.g. the addition of 5 selective nucleotides would reduce the percentage of false homology to 6.25% for non-homologous bands). This can be achieved for either primer, and so this process can be extended until a fragment is ultimately sequenced. While my results are conservative, for the congeners analyzed here (within *Onopordum* 0% cpDNA divergence and within *Cynara* 0.038% cpDNA divergence) many shared fragments were identified (19 between *Onopordum* spp. and 13 between *Cynara* spp., Table 3.2) but only 5% for *Onopordum* and 0% for *Cynara* were non-homologous. Therefore, as the frequency of such false homology could be expected to be three times less than that actually observed (75% vs. 25%), it is unlikely that many of these shared fragments are falsely homologous. This suggests that studies of phylogeny with AFLPs are best suited to closely related taxa.

While this PCR test for size homoplasy is conservative, it provides an extremely simple, and inexpensive method for detecting size homoplasy. Such a strategy should work equally for other multilocus techniques such as RAPDs. As most of the expense of AFLPs is incurred prior to the final PCR stage, the additional expense of running representative samples with the extra 4 or even 16 combinations of primers is not prohibitive. It is important to recognise however that this test is not a test of homology. Southern blotting remains a superior technique in that it can detect the differential

migration of alleles of the same locus and duplicated but separate genes, whereas the power of this test is in the identification of size homoplasy only. Nevertheless, due to its simplicity, this is a valuable technique to assess the likely bias of size homoplasy, or taxonomic informativeness with a given multilocus marker system in taxa of interest.

Chapter Four *An assessment of cpDNA sequence and AFLPs for phylogenetic analyses of the Carduinae thistles* (Asteraceae: *Cardueae)*

In this study I assess the utility of various molecular techniques in constructing phylogenies for some of the weedy Carduinae thistles. Using cpDNA sequence from the intron and inter-genic spacer of *trn*L-*trn*F and the spacer of *psb*A-*trn*H as well as nuclear AFLPs, I show that analyses from different sources provided contradictory phylogenies and also contradicted a previous morphological analysis. However, combining these data through different weighting schemes produced trees that were relatively resolved, with considerable support for many relationships. Sequence data from cpDNA provided limited and conflicting evolutionary signals. For AFLPs, tests for non-homology revealed that size homoplasy introduced a significant degree of bias at higher taxonomic levels but produced additional resolution at lower taxonomic levels. Where phylogeny is not well understood for particular groups, the creation and incorporation of large molecular data sets can easily be achieved. The increased phylogenetic resolution can provide additional information to help inform practitioners of biological control of thistles of the patterns of host-specialisation in target weeds.

Nucleotide sequence, especially derived from the chloroplast is an established technique, popular amongst molecular systematists because its interpretation (including homology) can be more straightforward than with other techniques. In addition, universal primers exist for many loci (providing a great number of potential characters), it is generally informative and powerful statistical methodologies have been developed for the data produced. Amplified fragment length polymorphism (AFLP) is a fingerprinting technique that combines elements of classic RFLP fingerprinting and anonymous PCR (Lin and Kuo 1995, Vos *et al.* 1995, O'Hanlon & Peakall 2000). Consequently, development time is minimal (Powell *et al.* 1996). While variable regions exist in the genome for sequencing,

their identification can be a long, difficult and costly process (Soltis *et al.* 1992, Whitkus *et al.* 1994, Hillis *et al.* 1996) particularly between close relatives where differentiation is moderate. Furthermore, studies sometimes require the analysis of many individuals or loci, and the labour intensive nature of sequencing makes this impracticable. In addition, because of its RFLP like nature, AFLPs are highly stringent and reproducible, overcoming some of the problems found with RAPDs (Stammers *et al.* 1995). Like other multi-locus profiling techniques, AFLP also provides the advantage of being a whole genome approach, removing the need for assuming the co-evolution of organisms and their individual genes. Despite these benefits, AFLPs like SSRs, RFLPs and RAPDs have a complex mutational process, and hence the data generated can difficult to interpret beyond identifying polymorphism (Powell *et al.*, 1996, Rieseberg 1996, Peakall *et al.* 1998, O'Hanlon & Peakall, 1999).

Highly variable multi-locus DNA techniques often produce co-migrating non-homologous bands (size homoplasy). In an analysis of RAPD fragments in hybridizing wild sunflowers, Rieseberg (1996) showed as many as 21% of 220 co-migrating fragment pairs were non-homologous. AFLPs have recently been applied to phylogenetic questions (Kardoulous *et al.* 1998, Huys *et al.* 1996). In such analyses, the utility of AFLPs has been discussed in a qualitative sense, regarding its ability to produce a resolved phylogeny of intuitive topology. Recently, the frequency of size homoplasy spanning different taxonomic boundaries has been quantified for AFLPs using a simple PCR test showing that homoplasy can be as high as 95% between genera but as low as 0% between congeners (O'Hanlon and Peakall, 2000; see also *Chapter Three*). As the present study

spans this range of taxonomic divergence, AFLPs alone may not be highly informative. However, it is possible that the combination of very conservative cpDNA sequence variation, the more rapidly evolving morphology, and the very rapid AFLP data may produce a more fully resolved phylogeny than any individual technique.

In this chapter, I examine the use of these molecular and morphological techniques to resolve problems in the phylogenetic relationships of the Carduinae thistles. A practical motivation underpinning phylogenetic studies in these weed species is that, as targets of biological control with high degrees of host-specialization, tests for host-specificity are required to ensure minimal risk to non-target organisms. Recent studies suggest strong links between the ability of a specialist agent to feed on non-preferred host plants and the evolutionary history of plant taxa involved (e.g. Futuyma *et al.* 1995, Kelley & Farrell 1998), and there has been an increasing interest amongst evolutionary ecologists in the subject of host-range and specialization by insect herbivores (see Jermy 1984, Jaenicke 1990, Courtney & Kibota 1990, Bernays & Chapman 1994). In particular, a number of papers have addressed the use of phylogenetic studies to understand how host-choice has evolved (see Mitter *et al.* 1991, Futuyma *et al.* 1995, Kopf *et al.* 1998, Kelley & Farrell 1998). Recent reports critical of biological control for causing non-target impacts (e.g. Simberloff & Stiling 1996, Louda *et al.* 1997) have focussed attention on this issue (see Wajnberg *et al.* 2000) and the testing methodologies needed to ensure that the host ranges of potential biological control agents are known (see Withers *et al.* 1999). The key, then, to adequate testing and minimizing risk is the availability of accurate information on phylogenetic relationships of host plants, and recent advance in molecular methods (e.g.

Nissen *et al.* 1995, O'Hanlon & Peakall 1999) can provide clearer pictures of this (Wagenitz 1997).

The Carduinae thistles are particularly appropriate taxa for such research for despite significant effort, relationships between these thistles remain ambiguous. Moreover, many taxa within this group have become important invasive weeds and are targets of biological control (see Julien & Griffiths 1998). Furthermore, it will remain difficult to understand the evolution of host-choice by potential biological control agents without an accurate estimate of host phylogeny. The specific objectives of this chapter are to (i) assess the utility of cpDNA sequence from two regions; trnL-trnF (inter-genic spacer and intron) and the psbA-trnH inter-genic spacer, and construct a phylogeny from these data, (ii) examine the utility of amplified fragment length polymorphism (AFLP) in reconstructing interrelationships between members of the Carduinae thistles, and (iii) to reconcile these results with each other, and with pre-existing morphological data.

Materials and Methods

Plant taxa and sampling

The Carduineae (Dittrich 1997) is a monophyletic subtribe of the tribe Cardueae (Asteraceae) comprising many of the widely known thistles. Based on morphological data (Petit 1997), the Carduineae itself appears to contain three monophyletic series; (i) the *Xeranthemum* related genera (Petit 1997), (ii) *Staehelina* (Petit 1997) and (iii) the genus rich "Carduineae-Centaurineae" *sensu* Dittrich (1977). This third series can be further divided into several monophyletic groups, two of which are; the relatives of *Carduus* and

the relatives of *Centaurea,* both highly speciose groups (Petit 1997). Here I focus on the relatives of *Carduus*.

Seed were imported from the Montpellier Botanic Institute and the Botanical Museum of Berlin. These, as well as additional seed collected by Briese D.T., Sheppard A., and O'Hanlon P. were germinated in quarantine at CSIRO, Canberra, Australia to obtain DNA and produce voucher samples where source maternal vouchers did not already exist (Table 4.1). Species were selected to replicate the genera within the *Carduus* clade of the Carduineae subtribe, described by Petit (1997), with outgroup genera belonging firstly to the tribe Echinopeae (*Echinops*), and secondly being basal to the "Carduineae-Centaurineae" series (*Jurinea*). As a preliminary study of phylogeny, and an assessment of the technqiues, each genus was represented by a single species except for *Cynara*, which was once considered paraphyletic (Wiklund 1992), and *Onopordum*. At least two samples per species were analyzed with cpDNA and AFLP techniques, and each sample was replicated to ensure that genetic data was reproducible, and that coding and scoring was accurate.

DNA methods

Total genomic DNA was extracted from seeds that were ground in liquid nitrogen using the procedure of Gilmore *et al.* (1992). DNA was precipitated in ethanol and resuspended in 50uL of low TE buffer. Two regions of the chloroplast were sequenced: *trn*L-*trn*F (inter-genic spacer and intron) and *psb*A-*trn*H (inter-genic spacer). These regions were selected as primers were readily available and the regions had previously been shown to

be informative within various *Asteraceae* (Bayer & Starr, 1998). For amplifying *trn*L-*trn*F, primers were synthesised for the sequences 5' CGAAATCGGTAGACGCTACG and 5' ATTTGAACTGGTGACACGAG as of Taberlet *et al.* (1991). I was unable to amplify *trn*L-*trn*F from *Jurinea*. For amplifying *psb*A-*trn*H, primers were provided by Randall Bayer (CANB) of sequence 5' GTTATGCATGAACGTAATGCTC and 5' CGCGCATGGTGGATTCACAAATC as of Sang, Crawford and Stuessy (1997). All reactions for both regions were identical. Approximately 40ng of DNA was used in a 40uL reaction containing 1 X PCR Buffer II (Perkin Elmer), 2.5mM of each dNTP, 8 pmol of each primer, 2.5U of Taq polymerase and 4.3mM of Mg^{2+}. Thermal regimes were identical to that used by Taberlet *et al.* (1991). After precipitation in EtOH (90uL) and ammonium acetate (10uL 3M), collection by centrifugation, and removal of the supernatant, the pellet was dried and resuspended in 20uL TE.

cpDNA sequencing

Sequencing reactions were performed with 20ng of template PCR product, amplified using ABI Prism "dye terminator" cycle sequencing chemistries (Perkin Elmer) according to the manufacturers instructions. Sequences were prepared for loading onto a 5% denaturing polyacrylamide gel set up in an ABI 377 automatic sequencer, according to instructions provided by ABI. Samples were electrophoresed and data collected using ABI collection software. Following electrophoresis and data collection, sequences were imported into Sequencher v 3.0 (GeneCodes) to assemble forward and reverse sequences. Sequence alignment was perfomed using Clustal V (Higgins *et al.* 1992) and required interpretation

of small (1-10 bp) insertions/deletions (indels). These indels were coded as "missing data" and were subsequently scored as binary vectors for presence/absence, and combined with the sequence data. Edited sequences have been submitted to GenBank (Table 4.1).

Table 4.1. Species of Carduinae and Echinopsidinae studied, sources of material, and GenBank accession of sequence data.

Species	Locality	Voucher	Genbank accession	
			trnL-F	psbA-trnH
Echinopsidinae				
Echinops sphaerocephalus L.	Alcaraz, Spain	not yet available	AF129830	AF129844
Carduinae				
Arctium lappa L.	Montpellier Herbarium, France	not yet available	AF129824	AF129838
Carduus nutans L.	CSIRO Entomology, Canberra, Australia	CANB not yet available	AF129825	AF129839
Cirsium vulgare (Savi) Ten.	CSIRO Entomology, Canberra, Australia	CANB not yet available	AF129826	AF129840
Cousinia hystrix C.A. Mey.	Goettingen Botanical Garden, Germany	CANB not yet available	AF129827	AF129841
Cynara cardunculus L.	Olympia, Greece	CANB JK not yet available	AF129828	AF129842
Cynara humilis L.	Albacete, Spain	CANB TT not yet available	AF129829	AF129843
Galactites tomentosa Moench	Montpellier Botanical Garden, France	not yet available	AF129831	AF129845
Jurinea humilis (Desf.) DC.	Montpellier Botanical Garden, France	not yet available		AF129846
Notobasis syriacus (L.) Cass.	Seed exchange B.G. Dijon.	Cultivated in BGBM, 158-07-96-70. 1.8.1996 Häffner (B).	AF129832	AF129847
Onopordum acanthium L.	Teruel, Spain	CANB TT not yet available	AF129833	AF129848
Picnomon acarna (L.) Cass.	Coatia, Insel Hvar, Seed exchange B.G. Marburg.	Cultivated in BGBM 068-32-87-14. 8/98 Häffner (B).	AF129834	AF129849
Ptilostemon afer (Jacq.) Greuter	Goettingen Botanical Garden, Germany	CANB not yet available	AF129835	AF129850
Silybum marianum (L.) Gaertn.	Corinth, Greece	CANB JK not yet available	AF129836	AF129851
Tyrimnus leucographus (L.) Cass.	Baillarguet, France	CANB not yet available	AF129837	AF129852

AFLP fingerprinting

AFLP reagents come in kit form and were purchased from Life Technologies. Fluorescently labeled primers were supplied by Perkin-Elmer. The procedure for generating AFLP fragments follows the instructions provided by the manufacturer and can be found in *Chapter Three.*

Identification of size homoplasy

Tests for size homoplasy between co-migrating AFLP fragments was conducted using a simple PCR-based approach (O'Hanlon and Peakall, 2000; see also *Chapter Three*) who applied the test to the same taxa studied here. Original selective PCR with MseI-CAC for AFLP was followed by separate PCR reactions using MseI-primers that were one nucleotide longer, and hence more selective. For the present study, I utilised this approach with each of my samples (M-CACA, M-CACC, M-CACG & M-CACT). This allowed the identification of the nucleotide at the fourth selective PCR position in all of the AFLP fragments. Where fragments of the same size had a different nucleotide in this position, it was taken as an indication of non-homology of fragments. This enabled the coding of the presence and absence of different but co-migrating loci that are usually indistinguishable and construct a phylogeny using this additional information.

Phylogenetic analyses

Firstly, the complete set of trnL-trnF and psbA-trnH sequences and coded indel matrices were subjected to a branch and bound TBR search with character states specified as

unordered and unweighted. Secondly, the AFLP matrix was analyzed in a similar fashion. Next, the subset of taxa studied here from Petit's (1997) morphological data was analyzed in the same fashion for direct comparison. Finally, as similar topologies were obtained from each source, all data was combined to give a total evidence analysis (see Bremer 1996) with two different weighting schemes. Firstly, I conducted analyses using the entire data set, with each character equally weighted. Secondly, as I was analysing three distinct data types with different numbers of characters and distributions of character states, I sought to weight them so that different data sets contributed equally to the overall analyses. To this end each character within each set was weighted so that the sum of weights of each character within each set equaled one (each of the 22 parsimony-informative morphological characters $W_{morph}=1/22=0.045$, $W_{trnL\text{-}trnF\ \&\ psbA\text{-}trnH}=0.042$, and $W_{AFLP}=0.011$).

Results

cpDNA variation

Of the 769 aligned total sites in *trn*L-*trn*F, 145 contained gaps in one or more taxa. For the 473 *psb*A-*trn*H sites, 56 contained gaps. While the trnL-trnF region contained a higher percentage of gaps (18.9% to 11.8%), *Echinops* contained a single but very large 152 bp gap in *trn*L-*trn*F, accounting for the difference. In total including binary indels, 1242 characters were scored, of which 72 were present only in a single taxon and only 22 were variable and present in more than one taxon (Table 4.*2)*.

Table 4.2 Polymorphic sites for aligned sequences of trnL-trnF and psbA-trnH regions of chloroplast DNA.

	trnL-trnF	psbA-trnH
positions[a]	111122223333344445555666777777777 indels[b]	11111111122222222222333333333334 indels[b]
	3369024634583358846674478238112346678	999936777888900235667889223344556891111111111122
Taxon	3402775983096740142782605169120521815123456789	023516057267219325344361896747238377012345678901
Echinops	ACCCA---CACACCAACAAAGGGGCGGATATCTGACA101001011	AAAAGGTCTGCGGCACGTGGTTACCTGCAAGAGCGT111100001110
Cousinia	ACCCGGCTCAATCTAAACAACGGGCAGATATCTGAC-111001011	TTTTGG-AAGCGGACCGGAGTCACATGAAGTAGCGT101010001010
Arctium	ACACGGCGCAATCCAAACAACGGGCAGATATATGGTG111001001	TTTTAGGAAGCGGAACGGAGTCACCTGAAGTAGTGT111100001010
Cynara card.	AGCCGGCTCAATTCACACAACGGGCAGATATCTGACG111001011	TTTTGAGAAGCGGAACGGAGTTACATGAAGTAGCGC111100001010
Cynara humi.	CTCCGGCTCAATTCACACCACGGGCAGATATCTGACG111001011	TTTTGAGAAGCGGAACGGAGTTACATGAAGTAGCGC111100001010
Ptilostemon	ACCCGGCTCAATCCTAACAACGAGCAG-TAGCTAACG111001010	TTTTGGGAAGCGAAACGGAGTTACATGATGTAGCGT111100001010
Onopordum	ACACGGCTCAATCCAAACAACGGGTAGCTATCTGACG111001011	TTTTGGCAAGTGGAACGGAGTTCCATTAAATAGCTT110100001010
Carduus	ACCCGGTTAAATCCAAACATCGGACAGATATCTGACG011011011	TTTTGGGAAGCGGAACGGAGTTACAGGAAGTAGCGT111100000010
Cirsium	ACCTGGTTAAATCCAAACAACGGACAGATATCTGACG011001011	TTTTGG-AAGCGGAACGGAGTTACAGTAAGTAGCGT101100000010
Galactites	ACACGGCTCAATCCAAACAACGGGCAAATATCAGACG111001011	TTTTAGAAAGCAGAACGGAGGTACATTATGTAGCGT111100000011
Notobasis	ACCCGGTTCAATCCAAACAACCGACAGATATCTGACG110100011	TTTTGGGAATCGGAACAGATTTACAGTAAGTTACGT111101010010
Picnomon	ACCCGGTTCCATCCAAACAACCGACAGATATCTGACG111001011	TTTTG-GAAGCGGAACGGAGTTACAGGAAGTAACGT011100000010
Silybum	ACCCGGTTCAATCCAAACAACGGACAGATATCTGACG011001011	TTTTGGGAAGCGGAACGGAGTTAAAGGAAGTAGCGT111100000010
Tyrimnus	ACCCGATTAAATCCAAACAACGGACAGAGTTCTGA--011101111	TTTTGGGCAGCGGAACGGAGTTACAGGAAGTAGCGT111100100010
Jurinea		TTTTGGGCAGCGGAAAGGAGTCACATTAAGTAGCGT111100001000
	* * * * * * * * *	** * * *** ** * * * *

[a] Dashes indicate gaps; asterisks indicate positions of cladistically informative nucleotides that are numbered consecutively from 5' to 3'.
[b] Italic numbers indicate insertions/deletions by presence (1) or absence (0) of sequence in indel region: 1. 38-41bp 2. 122-134bp 3. 304bp 4. 386bp 5. 468-475bp 6. 487bp 7. 493bp 8. 608-616bp 9. 688-715 bp, 10. 159-169bp 11. 169bp 12. 172bp 13. 207bp 14. 212-217bp 15. 265-270bp 16. 298-312bp 17. 357-358bp 18. 393 bp 19. 394 20. 410 21. 423-433.

AFLP polymorphism

Each sample generated 30.5 (s.e.=1.86) fragments within the size range of 75-450bp. With one AFLP primer pair, a total of 487 fragments were identified, all of which were polymorphic. This is an extremely high number, but is understandable given the taxonomic breadth of this group, the fine-scale analysis of fragment size, and the inherent degree of polymorphism obtainable with AFLPs.

Many of the fragments produced were unique to genera (396), with the number of shared fragments being much less (91). In the context of AFLPs, these genera are extremely divergent and similarities between them are relatively difficult to find. Indeed, the similarities that were found were generally restricted to two or three genera with very few polymorphisms being shared between more taxa. Overall, of the 91 fragments shared between samples within this group, 53% of them were shown to be non-homologous (Table 4.3). While such a figure is extremely high, the frequency of size homoplasy was related to the taxonomic divergence between the samples. Between congeners, an average of only 2.5% size homoplasy was observed, while between species from different subtribes a high of 88% homoplasy was found (Table 4.3). Congeners shared many homologous fragments, but between taxa of different levels of divergence, there was no clear pattern to the number of co-migrating fragments (Table 4.2).

Table 4.3. Comparisons of the total number of shared AFLP fragments (lower triangular matrix), and the number of non-homologous but co-migrating fragments (upper triangular matrix) within the Carduineae. Lighter shaded cells represent comparisons across subtribes and darker shaded cells represent congeneric comparisons. (Same as Table 3.2)

	Ech	*Jur*	*Cou*	*Arc*	*Not*	*Pic*	*Tyr*	*Gal*	*Sil*	*Car*	*Cir*	*Cy. hu*	*Cy. ca*	*O. ac*	*O. il*	*Pti*
Echinops	-	0	2	4	0	4	1	3	0	2	1	0	0	0	1	0
Jurinea	1	-	1	0	2	4	2	4	2	1	4	2	4	1	3	2
Cousinia	3	2	-	2	1	3	5	2	0	1	0	2	3	1	2	2
Arctium	7	4	3	-	2	2	2	0	3	2	6	2	4	1	2	5
Notobasis	0	2	2	4	-	1	0	2	4	1	6	3	4	0	2	4
Picnomon	4	4	3	5	6	-	3	4	4	2	5	3	4	3	4	2
Tyrimnus	1	6	7	3	0	3	-	1	1	2	1	2	0	1	2	0
Galactites	4	5	4	1	2	4	2	-	1	2	0	2	4	2	2	0
Silybum	0	3	0	3	5	9	3	4	-	2	2	1	2	1	0	1
Carduus	2	3	1	3	2	3	8	3	4	-	4	1	0	0	3	2
Cirsium	2	5	3	9	7	9	3	0	7	5	-	2	5	2	3	3
Cy. humilis	1	3	2	3	4	3	2	2	1	1	3	-	0	0	2	0
Cy. cardunculus	1	4	4	5	6	4	1	5	2	0	8	15	-	1	2	2
O. acanthium	1	3	3	4	1	5	4	2	1	1	4	1	3	-	1	0
O. illyricum	3	5	4	3	2	4	7	2	0	4	5	2	5	19	-	3
Ptilostemon	2	5	3	7	4	4	0	0	1	2	4	3	4	3	4	-

Morphological phylogeny

Morphological analyses have been described in detail by Petit (1996). Re-analysis here produced a single shortest tree (length 44, RC 0.92, HI) with no topological changes, and allowed an assessment of the strength of individual branches not previously provided. All nodes showed moderate bootstrap support, except for that placing *Silybum* with *Galactites* and *Tyrimnus* (Fig. 4.2).

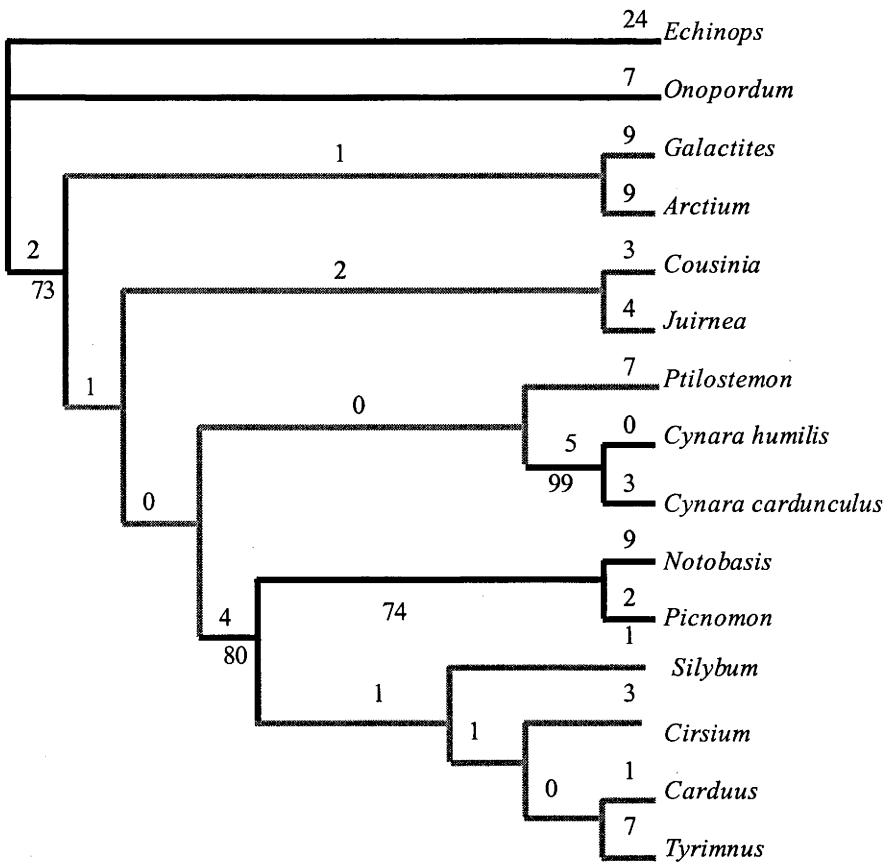


Figure 4.1 Single most parsimonious tree for cpDNA sequence data (length 108, CI 0.2433, RC 0.7312). Numbers on branches represent bootstrap values and those above the branches are consensus tree values. Grey nodes represent bootstrap support of <50%.

cpDNA sequence phylogeny

The 22 cladistically informative characters produced a single most parsimious tree (length 108, RC 0.73, HI). As expected the two *Cynara* species formed a monophyletic clade (bootstrap 99%, T-PTP 100%) as did *Picnomon* and *Notobasis* (bootstrap 74%, T-PTP 100%, Fig. 4.1). In addition, a well supported clade comprising *Notobasis*, *Cirsium*, *Picnomon*, *Silybum*, *Carduus*, and *Tyrimnus* (bootstrap 80%, T-PTP 100%) was formed (Fig. 4.1). While many unexpected relationships were uncovered in this analysis, they were supported by bootstrap values of less than 50%.

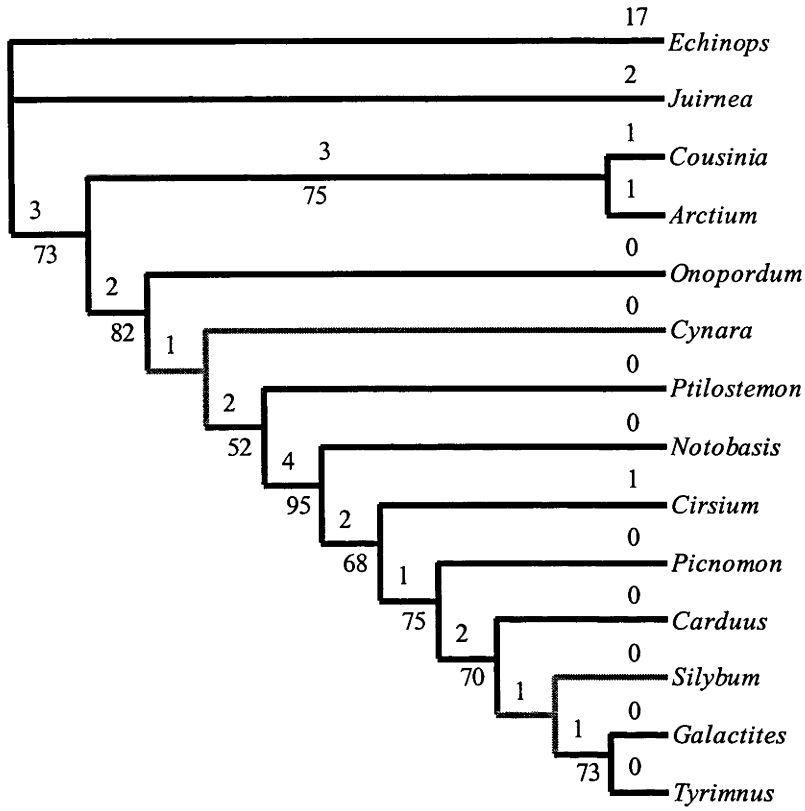


Figure 4.2 Single most parsimonious tree for morphological data (length 44, HI 0.1923, RC 0.8112). Numbers on branches represent bootstrap values and those above the branches are consensus tree values. Grey nodes represent bootstrap support of <50%.

AFLP phylogeny

For primers MseI-CACx, the 91 cladistically informative characters produced 51 shortest trees (length 222, RC 0.3542, HI 0.3018, Fig. 4.3). A strict consensus tree and bootstrap analysis failed to support most of the nodes identified, except for those supporting pairs of closely related genera.

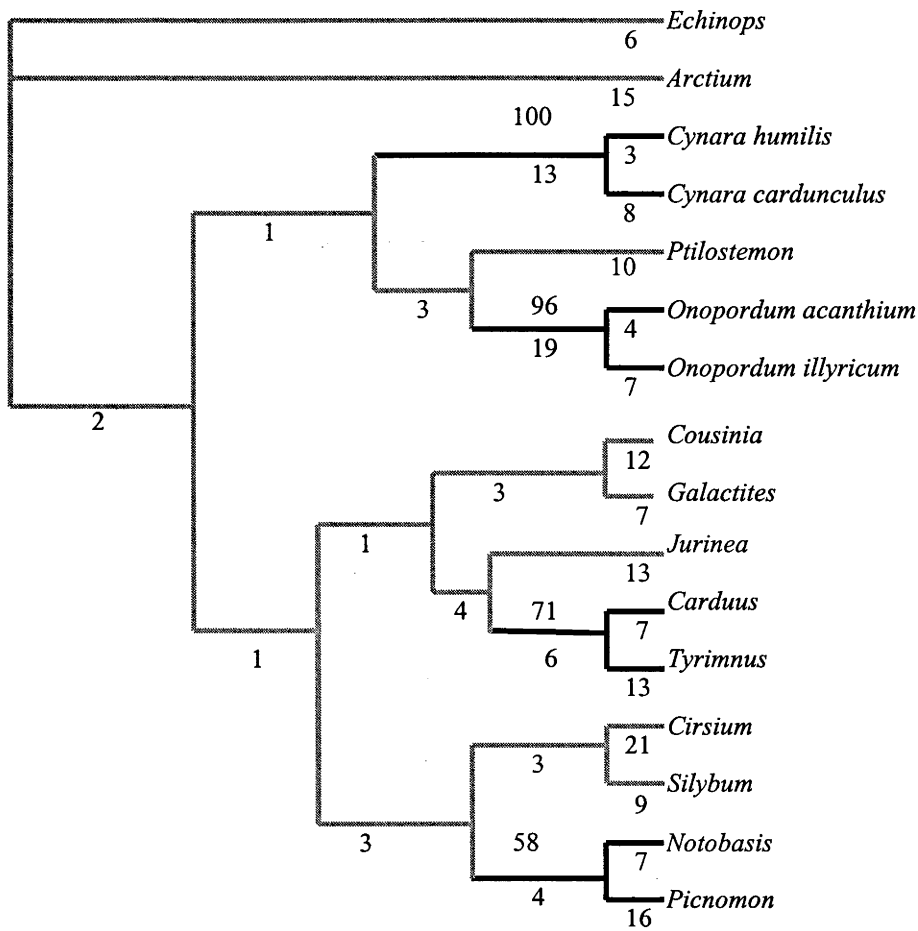


Figure 4.3 Bootstrap 50% consensus of 9 equally most parsimonious trees derived from AFLP data (length 221, HI 0.3122, RC 0.3255). Numbers on branches represent bootstrap values and those above the branches are consensus tree values. Grey nodes represent bootstrap support of <50%.

Reconciliation between data sets

When the three different data types were combined, a total of 137 parsimony informative characters were obtained, producing a single shortest tree (Fig. 4.4, length 431, RC 0.4371, HI 0.6612). However, when characters were weighted so that the three data sets contributed equally, three shortest trees were produced (results not shown).

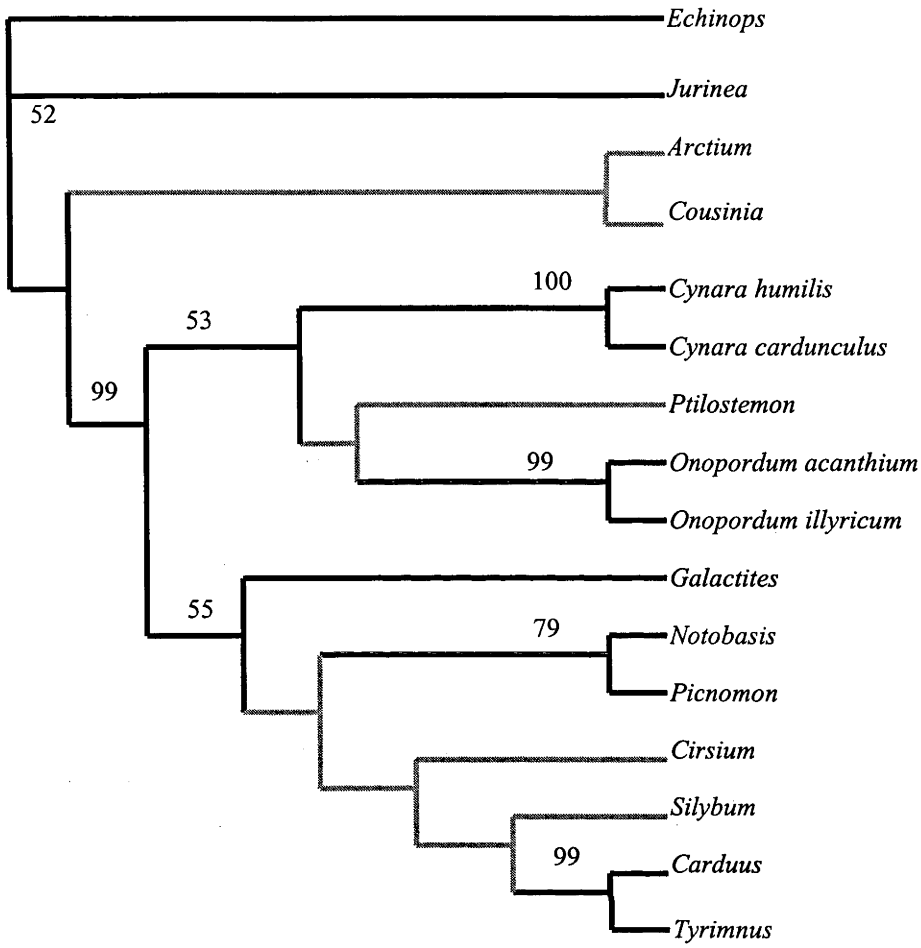


Figure 4.4 Single most parsimonious tree (length 431, HI 0.6612, RC 0.4371) of Carduinae thistles generated from a combination of AFLP variation, sequences of trnL-trnF and psbA-trnH, and morphological variation. Data sets combined in an unweighted fashion. Numbers above branches represent bootstrap values. Grey nodes represent bootstrap support of <50%.

Topological changes occurred only rarely when different weighting schemes were employed. However, bootstrap support for individual nodes often was substantially altered. The ingroup was supported by minimal length trees from both weighting schemes, and by a moderate bootstrap value when characters were weighted (64%). However, when all characters were equally weighted, the node separating *Cousinia* and *Arctium* from the outgroup genera *Echinops* and *Jurinea* was only poorly supported by bootstrap analysis (52%).

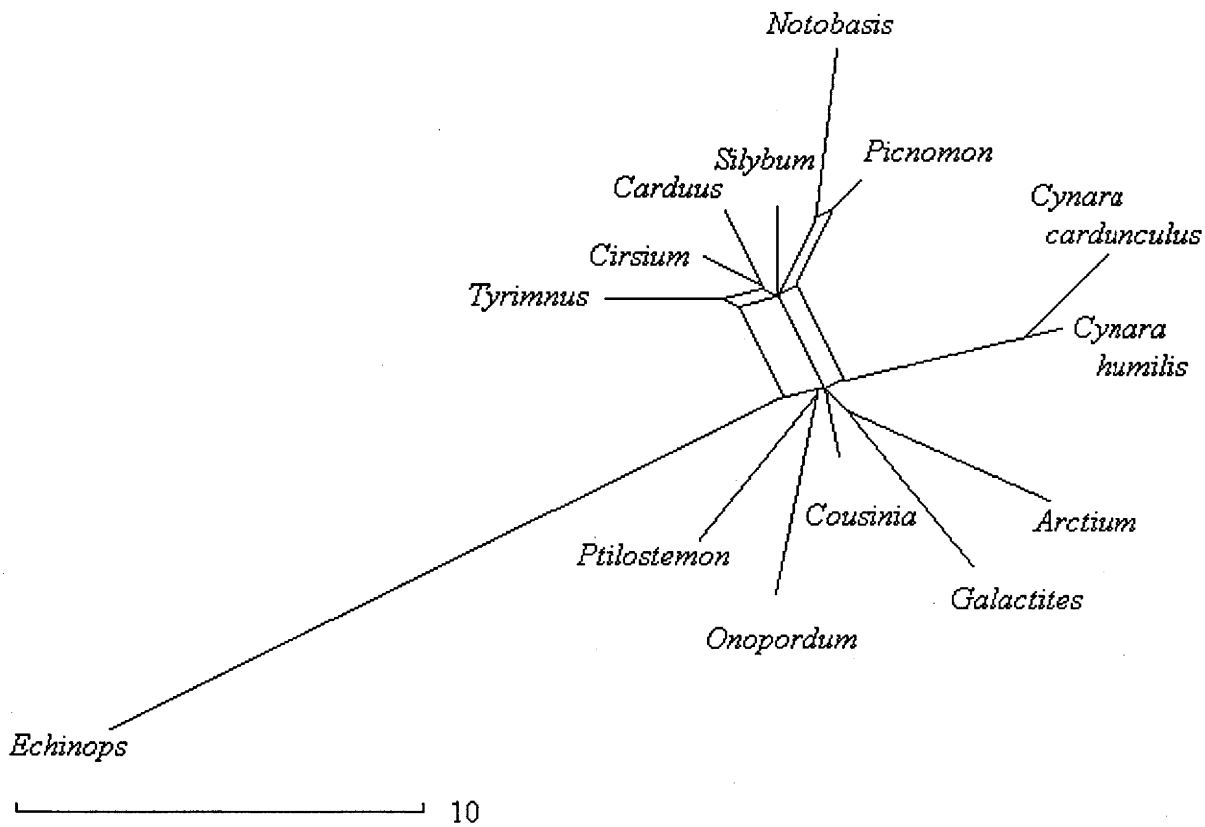


Figure 4.5 Split decomposition of chloroplast sequence data (Fit=81.5). Groups of parallel lines form splits. Clusters can be visualised by removing sets of parallel lines. The strength of the split is indicated by the length of the lines. In some circumstances sets of weakly compatible splits can be seen with individual genera being associated with different clades.

The positions of the *Cynara* species, the *Onopordum* species and *Ptilostemon* are equivocal for both weighting schemes (Fig. 4.4 & 4.5.). When AFLP data were included, bootstrap analysis of the unweighted data gave moderate support to a clade comprising *Cynara* and *Ptilostemon*. However, this value was very low. These three genera formed a polytomy with a clade including all the remaining genera (bootstrap weighted 89%, unweighted 99%). Within this group, a clade including *Carduus* and *Tyrimnus* was always formed, regardless of weighting scheme. Similar strong support for a clade including *Notobasis* and *Picnomon* was also found.

Discussion

Chloroplast variation

Considering the taxonomic breadth of this study, surprisingly low levels of cladistically informative cpDNA variation was revealed. A total of 94 mutations were observed, but only 22 were cladistically informative with the remaining 72 being unique to a single taxon. While increased sampling within genera would have reduced this ratio, unless genera were themselves highly polymorphic (not found within *Cynara* or *Onopordum*), synapomorphy between genera would remain at the same low level. While it is difficult to compare across taxa of different age, it is interesting to note the results of two major studies. In a more taxonomically restricted study of *Paeonia*, Sang *et al.* (1997) revealed a total of 35 mutations in *trn*H-*psb*A with 22 synapomorphs. When compared with the 12 synapomorphs observed here, the low chloroplast diversity in the Carduinae is revealed. A similar pattern was also observed for the *trn*L-*trn*F region, despite the fact that in addition to Sang *et al.* (1997), I sequenced the two *trn*L exons and the intron (*Paeonia*: nine synapomorphs; Carduinae: eight synapomorphs). Furthermore, using the *trn*L intron sequences, Gielly and Taberlet (1996) identified substantial nucleotide divergence between members of the same genus (*Gentiana*). Such a result obtained within the Carduinae indicates a relatively slow cpDNA evolution in the Carduinae relative to their morphological evolution. Therefore, many chloroplast loci may be required to obtain sufficient evolutionary signal for robust phylogenetic analysis.

AFLP polymorphism

The largest contributor to AFLP polymorphism were fragments unique to genera, indicating that the upper bounds of taxonomic divergence for resolving relationships was been exceeded

in this study. This has been confirmed by a PCR test for size homoplasy (O'Hanlon and Peakall, 2000) for the same genera studies here. In that study, up to 83% of co-migrating fragments were shown to be non-homologous between taxon-pairs, giving a false appearance of genetic similarity. Nevertheless, O'Hanlon & Peakall (2000) also identified a relationship between taxonomic divergence and false homology. Many shared and homologous fragments were identified between congeners, indicating that AFLPs are extremely useful, but that their utility is restricted to comparative studies between closely related taxa. That AFLPs are unsuitable at higher taxonomic levels is also reflected in the phylogeny produced with AFLP data alone, where only terminal taxa formed robust clades (Figure 4.3). This is not only a consequence of size homoplasy, but also of the limited numbers of shared fragments between taxa. Unlike both the cpDNA sequence and morphological variation, most shared fragments were relatively rare across the taxa, present in only one or two genera. While fragments were usually only shared between congeners and other close genera, increasing taxonomic divergence did not always result in fewer shared fragments, as predicted by Mueller and Ullrich (1999). For example, the number of shared fragments was sometimes relatively high between members from other subtribes (Table 4.3). Furthermore, the difficulty in predicting the number of shared fragments is illustrated by the fact that the *Onopordum* species had more than twice as many shared fragments as were shared between the distantly related *Cousinia* and themselves. In addition, the two *Cynara* species shared 13 fragments but the relatively distantly related *Cirsium* and *Cousinia* shared a relatively high nine fragments. Even though relatively little informative cpDNA sequence variation was revealed in this group, AFLPs evolve too rapidly to be useful for the same taxa.

Reconciliation between cpDNA sequence, AFLP and morphological data

Whether an unweighted or weighted scheme was employed did not markedly alter the final tree topology. However, the support from bootstrap analyses for individual clades did change, particularly at the deeper nodes. This was due to the high number of AFLP characters, which during bootstrap re-sampling of the data matrix, would have swamped the evolutionary signal provided by the fewer morphological and chloroplast data. Therefore, for studies combining the potentially numerous AFLP loci with fewer but more informative characters derived elsewhere, the differential weighting of data types may be necessary.

When molecular data was combined with morphological data, the overall topology was retained, with *Echinops* and *Jurinea* as outgroup, and *Arctium* and *Cousinia* separating relatively early from the rest of the ingroup. Additionally for the remaining genera, one major clade was formed by *Galactites*, *Notobasis*, *Picnomon*, *Cirsium*, *Silybum*, *Carduus* and *Tyrimnus*, and another group, *Onopordum*, *Ptilostemon* and *Cynara*, though this group was a loose one. The additional data provided by molecular techniques substantially altered the interrelationships within the major clade, with virtually no agreement. Both cpDNA and AFLP data placed *Notobasis* and *Picnomon* together, and this pattern was retained when combined with morphological data; and similarly for *Carduus* and *Tyrimnus*. Finally, the position of *Galactites* is the most flexible depending on the source of the data. AFLP and cpDNA sequence placed *Galactites* first with *Jurinea* and secondly with *Arctium*, and in neither case was the relationship well supported. Morphological data placed *Galactites* with *Tyrimnus* so it was not surprising that the combination of data produced a tree with *Galactites* at the base of the major clade containing *Notobasis*, *Picnomon*, *Cirsium*, *Silybum*, *Carduus* and *Tyrimnus*.

Implications for host-choice studies

For shifts in host affiliation, demographic constraints, such as relative host abundance, population genetic constraints such as restricted or ongoing gene flow, and ecological constraints such as synchrony between life cycles of potential hosts and insects, may be nested within the phylogenetic constraints of both the agents and the hosts (Briese 1996). As new markers such as AFLPs, which are most useful for phylogenetic studies at lower taxonomic levels, may be used in conjunction with more established methodologies such as sequencing, resolution of otherwise difficult groups may be more easily attainable than previously, assisting in studies of host-specialization and host-shifts. However, to understand fine scale evolutionary processes, robust topologies are necessary as errors may be magnified when comparing phylogenies. The data obtained in this study are not likely to prove to be suitable for such phylogenetic comparisons between hosts and pests.

As the thistles are an important group targeted for biological control by insects, there is a high potential to conduct studies relating phylogeny to host-specialization and host-shifts within this group. While this study is not conclusive, the improved understanding of phylogenetic patterns within the Carduinae thistles derived here provides additional background information with which to examine these questions. Part of this process may be the design of host-specificity tests based on the host phylogeny, enabling biological control workers to define host-specificity with more assurance. However, further refinement of the phylogeny within the Carduinae and Centaurine thistles is first required to test specific hypotheses.

Chapter Five *Recolonization of Europe from hybrid refugia by Onopordum thistles: a phylogeographic analysis*

Episodes of glaciation, refugial isolation and recolonization have been critical in the evolution of natural plant populations, and the shaping of contemporary patterns of plant population genetic structure. Here I show that European *Onopordum* thistles, which are a mostly Mediterranean group, display several features relevant to population structure common with other Mediterranean plant taxa. Using AFLP and cpDNA sequence data I demonstrate (i) that there is a strong genetic divide between species occurring in the Iberian Peninsula and the Aegean region and (ii) that the species poor northern group forms a distinct lineage. While such a pattern has been repeated in other taxa, it is highly incongruent with the morphological variation in *Onopordum*. As many morphological characters are uncorrelated with the phylogeny (as well as with morphological classification), it is unlikely that they all evolved separately on several occasions. Rather, it is more probable that this range of traits was present prior to speciation and that the isolation of the Aegean and Iberian taxa preceded the most recent major European speciation in this group. During glacial maxima where refugia were restricted to the extreme south, widespread hybridization may have resulted in highly polymorphic ancestors which, when re-invading Europe during glacial retreat, speciated with the sorting of morphological traits.

Episodes of glaciation, refugial isolation and recolonization have been shown to have been critical in the evolution of natural plant populations, and in shaping contemporary patterns of plant population genetic structure (Hewitt 1996; Thompson 1999; Taberlet *et al.* 1998). The importance of such processes is particularly evident in numerous examples of plant species from Europe. Hewitt (1996) has examined the effects of the ice ages on European species in detail. At the onset of glaciation cycles, many plant taxa underwent contractions into lower and warmer latitudes. At glacial maxima, many taxa existed only in southern refugia, bounded to the north by tundra or ice sheets and to the south by the Mediterranean Sea (Hewitt 1996). As warming began and glaciers retreated, a northward expansion from refugia, involving spreading

from the leading edge, ensued (Hewitt 1996). Typically, contemporary patterns of population genetic structure of plant species and other relatively sedentary groups such as non-migrating insects reveals a loss of genetic diversity in northern populations due to the bottlenecks that such a process entailed (Thompson 1999). The last northward expansion of European trees, which occurred from 13ky BP, is well documented from the pollen record (Huntley and Birks, 1983), and exhibited a very rapid rate of spread of up to 2000m/yr (Bennett 1986). Several species are believed to have recolonized the main land mass of Europe from refugia from the Iberian Peninsula, Southern Italy, the Balkans and the Aegean regions, which could have each acted as refugia for coastal Mediterranean species (Taberlet *et al.* 1998). In fact, as the tundra area probably extended as far south as Southern France, and sea levels were lower it is likely that suitable habitat for many coastal species only occurred off the current coastline, in the present area of the Mediterranean Sea (Thompson 1999). Such strong restrictions may account for the many species with disjunct distributions, the numerous local endemics and the strong spatial patterns to population genetic structure present in the European flora (Thompson 1999).

Ecological factors associated with recolonization after glaciation have also produced structure. Hewitt (1996) describes "suture zones" where taxa moving out of refugia have come into contact. At such places hybrid zones have acted as barriers, reinforcing speciation. In the cricket *Chorthippus parallelus*, a relatively narrow cline has been produced by hybridization in the Pyrenees, separating French and Spanish genomes (Barton and Hewitt 1989). As there has been little introgression through this hybrid zone for perhaps 9000 years, it has acted to protect the integrity of the two genomes during the warmer interglacial period (Hewitt 1996). There are several other

examples of hybrid zones apparently acting to reinforce species integrity (see Hewitt 1996; Taberlet 1999). However, hybridization can also act to homogenize once-separated populations, and eliminate population structure. Comes & Abbott (1999) showed such homogenization between the chloroplasts of two hybridizing *Senecio* species. Analagous to, but outside the context of natural colonization patterns are the examples of hybridization between native and alien species, leading to the endangerment of the native species (Levin 1996) through the breakdown in species boundaries. Such processes could eliminate historical biogeographic patterns and remove any population structuring between recolonizing taxa.

The nature of a hybrid swarm is such that it can contain more genetic diversity than either parental species alone (Emms and Arnold 1997; Rieseberg and Carney 1998). Refugial populations containing hybrids may therefore contain more potentially adaptive traits of value during periods of environmental change, or during recolonization (see Thompson 1999).

There is no general rule on the influence of hybridization on the evolution of plant populations (Arnold and Hodges1995; Rieseberg and Carney 1998; Fuertes Aguilar *et al.* 1999), but it is clearly an important evolutionary process in the context of recolonization, either as a reinforcing mechanism leading to speciation or as a mechanism allowing the transfer of adaptive traits and the production of evolutionary novelty.

For these reasons, hybridizing Mediterranean taxa are of special interest in the study of the evolution of natural plant populations and speciation. The *Onopordum* genus of thistles (*Asteraceae*: Cardueae) has a Eurasian origin and can be found from central Asia to western Europe and North Africa, (as well as having been introduced into

Australia and North and South America). However, all but one of the species in Europe (*O. acanthium*) have a characteristically Mediterranean distribution. Thus, during the glacial maxima *Onopordum* was probably restricted to small patches at the extreme southern tip of the continent. The current distribution appears to reflect such a history. In Europe, there are two centres of species richness; the Aegean region, and the Iberian Peninsula (Franco 1976). Between these centres, *Onopordum* is relatively rare, often occurring as a ruderal species. However it is commonly found on the Mediterranean islands. The distribution of these species in Europe is therefore very similar to several other Mediterranean species, with a strong geographic divide between species found in the Aegean and those in Spain. In addition, only one species occurs in higher latitudes, reminiscent of the limited diversity observed in the northern parts of the distribution for many taxa (Thompson 1999).

Over 100 species of *Onopordum* have been described (Dress 1966), although many described species are synonymous and the actual number is probably closer to 50 (Michael 1996). In Europe, there are thirteen species recognised, six of which have subspecies (Franco 1976). Morphologically, the genus is notably diverse, with species showing gross differences in leaf colour and morphology, floret colour, flower head size and shape, plant architecture and involucral bract size and shape. Life-history traits, such as flowering phenology and perenniality also vary between species.

Despite this abundance of variation, there has been no attempt to construct a phylogeny for the genus, with most effort being dedicated to the taxonomic delineation of species boundaries (Franco 1976; Michael 1996; P. Hein, pers. comm. 1999). This is probably due to the high levels of polymorphism within species. Further complicating these difficulties are observations of hybrids between many of

these species in Spain (Gonzalez Sierra *et al.* 1992), Italy (D. Briese, pers. obs.), Greece (A. Sheppard pers. comm. 1999), Turkey (P. Hein pers. comm. 1999) and Australia (O'Hanlon *et al.* 1999). In the latter case, widespread hybridization has occurred following invasion and colonisation of a new habitat by different *Onopordum* taxa (O'Hanlon *et al.* 1999). Not only are hybrids believed to occur between species within the same subgeneric classifications of Franco (1976), but also between species at morphological extremes within the genus (Gonzalez Sierra *et al.* 1992). The genus thus appears to have weak internal reproductive barriers. *Onopordum* is therefore a genus that provides opportunities to study not only the consequences of glaciation and recolonization for speciation, but also the role that hybridization has played in shaping current genetic patterns.

My specific objectives for this chapter are to; (i) construct a molecular phylogeny of European *Onopordum* thistles, (ii) to determine whether the resultant phylogeny was geographically structured, (iii) to relate any such structure to known glacial refugia, and (iv) to draw any appropriate evolutionary implications. As a rsult, evidence was found that hybridization within isolated refugia produced phenotypically diverse founders for the post-glacial recolonization of Europe, and that subsequent recolonization resulted in the sorting of adaptive traits, giving a distribution of morphological characters incongruous with the molecular phylogeny.

Materials and Methods

Sampling

Seed was imported and germinated in quarantine at the CSIRO Black mountain Quarantine Facility, Canberra, Australia. Sampling was conducted to represent most of the European species of *Onopordum*, with outgroup species (*Cynara humilis* and *Cynara cardunculus*) belonging to a sister genus. Seed was collected and germinated from different sites and subspecies as shown in table 5.1. These sites were considered to be 'pure' representatives. The species *O. laconium* and *O. caulescens* are both Balkan species and are unfortunately omitted from the present analysis due to their absence from accesible herbaria and financial limitations.

Table 5.1. Carduineae species examined according to the classification of Franco (1976) and localities from which seed was obtained.

Subgenus SECTION	Species	Location
Acaulon.	*O. acaulon acaulon* L.	Sierra Segura, Spain
	O. acaulon acaulon L.	Zaragoza, Spain
	O. acaulon uniflorum (Cav.)	Llo-E. Pyrenees, France
Onopordum. ONOPORDUM	*O. acanthium* L.	Cuenca, Spain
	O. acanthium L.	Florina, Greece
	O. acanthium L.	Tobia, Italy
	O. acanthium L.	Eyguieres, France
ECHINATA Franco.	*O. macracanthum* Schousboe	Cordoba, Spain
	O.argolicum Boiss.	Tripoli, Greece
	O. corymbosum Willk.	Teruel, Spain
	O. tauricum Willd.	Lake Volvi, Greece
RECURVATA Franco.	*O. bracteatum bracteatum* Boiss. & Heldr. In Boiss.	Kastoria, Greece
	O. bracteatum ilex (Janka) Franco	Lamia, Greece
	O. bracteatum myriacanthum (Boiss.) Franco	Argos, Greece
	O. bracteatum myriacanthum (Boiss.) Franco	Epidavros, Greece
	O. illyricum L. *illyricum*	Tobia, Italy
	O. illyricum L. *illyricum*	Bonifacio, Corsica
	O. illyricum cardunculus (Boiss.) Franco	Igoumenitsa, Greece
	O. illyricum cardunculus (Boiss.) Franco	Ionnina, Greece
ERECTA Rouy.	*O. nervosum* Boiss.	Cuenca, Spain
Putative hybrids	*O. acanthium X O. illyricum*	Tobia, Italy
	O. acanthium X O. illyricum	Eyguieres, France
	O. corymbosum X O. acanthium	Teruel, Spain
	O. nervosum X O. acanthium	Cuenca, Spain
	O. bracteatum br X O. acanthium	Lefkonas, Greece

DNA extraction

Total genomic DNA was extracted from seeds that had been ground in liquid nitrogen using a modification of the procedure of Gilmore *et al.* (1992). A single chloroform treatment was followed by purification according to the diatomite method of Gilmore *et al.* (1992) to remove remaining protein, polysaccharides and RNA. DNA was precipitated in ethanol and resuspended in 50uL of low TE (10mM Tris-HCl (pH 8.0), 0.1 mM EDTA) buffer. DNA was used for sequencing and AFLP analysis. Replication of DNA extractions was performed using different plants and for each extraction, sequencing and AFLP analyses were duplicated.

Chloroplast sequencing and AFLP fingerprinting

The chloroplast region, *trn*L-*trn*F is a highly variable region consisting of two genes, and two introns usually being about 900bp in total length. Sequencing reactions, purification, and sequence alignment and editing were performed as of Chapter Four. Sequence obtained was checked against GenBank, showing close similarity with other members of *Asteraceae.*

Recent work has shown that at low levels of taxonomic divergence AFLPs are useful for phylogenetic reconstruction (Chapter Four) and do not suffer strong bias introduced by size homoplasy. Using a PCR-based test for size homoplasy, of the 19 fragments shared between two *Onopordum* species (*O. illyricum* and *O. acanthium*) and 13 fragments between *Cynara* species (*C. cardunculus* and *C. humilis*) only 1 and 0 fragments respectively were found to be non-homologous.

Digestion, annealing and pre-selective amplification steps of AFLP fingerprinting were performed as of Chapter Four. Selective PCR was performed by combining 2.5uL of each diluted, pre-selectively amplified sample with 3.95 uL MilliQ water, 0.25uL of blue labeled EcoRI-aca primer, 2.25uL of MseI-CTA, 1.0uL 10X PCR buffer, and 0.25 units Taq polymerase. PCR commenced with 94°C for 30 seconds, 65°C for 30 seconds, and 72°C for 60 seconds. In subsequent cycles, annealing temperature was reduced by 1°C per cycle until it reached 56°C, followed by 23 cycles at 56°C. The final PCR products were electrophoresed in 5% denaturing acrylamide in an ABI Prism 377 sequencer where profiles were recorded digitally. Fragments were accurately sized by the inclusion of an internal size standard. Digital profiles were visualized and scored with the aid of ABI GeneScan 2.5.

Phylogenetic analysis

As hybridization is known to occur between members of this group, split decomposition analysis (Huson 1998) was performed. It has been shown to be useful in reconstructing histories in other hybridizing plant groups (Hollingsworth *et al.* 1999). Split decomposition is non-hierarchic technique in that it identifies sets of splits within the data and produces a network of weakly incompatible splits. The strength of the splits can be shown visually using a network, and hence provides a set of evolutionary hypotheses. Where data is tree-like, the network is also a tree. However, where data contains conflicting hierarchic signals, such as will be the case after reticulation, alternate hypotheses are provided. Split decomposition was conducted using SplitsTree 3.1 (Huson 1999) with 1000 bootstrap replicates and the Hamming distance option. In addition, to aid visualization of non-hierarchical structures within the data, I performed non-metric multidimensional scaling (NMDS)

on pairwise squared Euclidean genetic distances between all individuals using SYN-TAX (Podani 1986). NMDS was performed 30 times to avoid reaching a local rather than global minima, and the analysis with the lowest stress value was used.

Cladistic and phenetic methods were used to construct hierarchical phylogenetic relationships between all samples, except for those identified from the field or from split decomposition as being of putative hybrid origin. An exhaustive search for the shortest tree was conducted, and bootstrap consensus trees were calculated for branch and bound parsimony searches, as well as for neighbor joining trees. All analyses were conducted using PAUP 4.0b (Swofford 1999). Using Franco's (1976) description of the species studied, I mapped the distribution of certain morphological characters onto my phylogenetic tree to examine the evolution of particular morphological characters.

Results

Chloroplast and AFLP polymorphism

The region was sequenced with no ambiguity for 800 base pairs for each species. Despite the fact that *Onopordum* species are morphologically and ecologically very diverse (10cm-300cm tall, annual-perennial), there was no variation detected in the inter-genic spacer and intron of this chloroplast region. Within other genera, chloroplast divergence has been identified (Mes and Thart 1994; Kita *et al.* 1995; Gielly and Taberlet 1996; Kim *et al.* 1996; Fujii *et al.* 1997; Maguire *et al.* 1997; Mes *et al.* 1997; Sang *et al.* 1997; Cros *et al.* 1998; Lanner 1998). While members of *Asteraceae* typically have little chloroplast divergence (Bayer and Starr 1998), that no variation was observed here was nevertheless surprising. In contrast, with a single

AFLP primer pair combination, a total of 160 polymorphic fragments were identified giving many markers for phylogenetic analysis. Of these, 27 were unique to *Cynara*, and 33 fragments were unique to individual species in *Onopordum*.

Phylogenetic analyses

Non-metric multidimensional scaling produced a two dimensional plot, with a stress value of 10.4%, from the original multidimensional genetic distance matrix (Fig. 5.1). This figure is relatively high, showing that for the number of samples, the distance matrix can only be represented to 89.6% accuracy in two dimensions. Within species, samples were closely clustered. Several species such as *O. illyricum* and *O. bracteatum*, as well as *O. acanthium* and *O. corymbosum* appeared closely related. A major divide between species found on the Iberian peninsula and around the Aegean is apparent. Several samples identified in the field putatively as of hybrid origin were found to be genetically intermediate between putative parental species. In addition, one sample identified as *O. bracteatum,* subsp. *bracteatum* appeared intermediate between *O. acanthium* and *O. bracteatum*. This sample was collected from a site containing both *O. acanthium* and *O. bracteatum*.

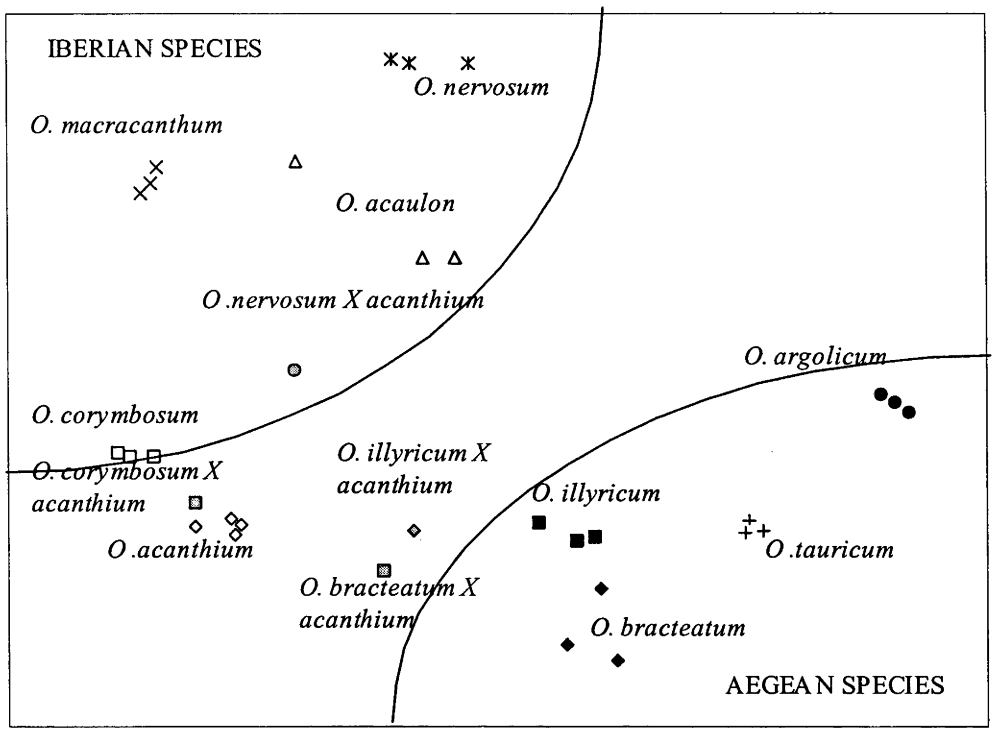


Figure 5.1 Two-dimensional non-metric multidimensional (NMDS) scatterplot of pairwise genetic distances between European members of *Onopordum* species (stress value=10.4%) showing non-hierarchical divide between Iberian and Aegean species (dotted lines). *O. acanthium* falls into neither of these groups and represents the only species found in Northern Europe. Several samples (gray symbols) represent putative hybrids collected where parental species co-occur.

Split decomposition analysis produced a non-hierarchical tree that was mostly concordant with NMDS analysis (Fig. 5.2). Species that have more than a single possible relationship with other species are connected by pairs of parallel lines with several species, rather than a single branch to its nearest neighbor. The removal of different pairs of parallel lines places individual taxa in "splits" with other sets of taxa with the length of the parallel lines, and bootstrap values indicating the strength of the split. All samples identified as putative hybrids from the field could be clustered compatibly with either of its putative parents, indicative of hybridization. Similar to the NMDS result, the sample identified as *O. bracteatum,* subsp. *bracteatum* formed compatible splits with both *O. acanthium* and *O. bracteatum* (Fig. 5.2). Split decomposition analysis failed to reveal strong hierarchical signal within AFLP genetic data, and most species appeared to radiate from a single point.

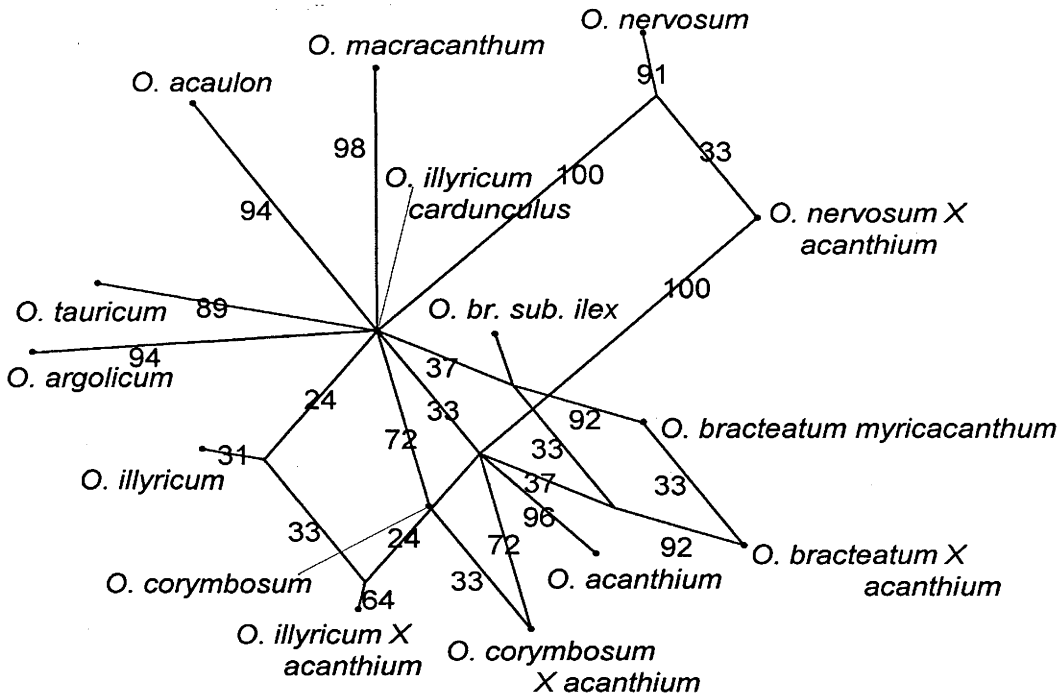
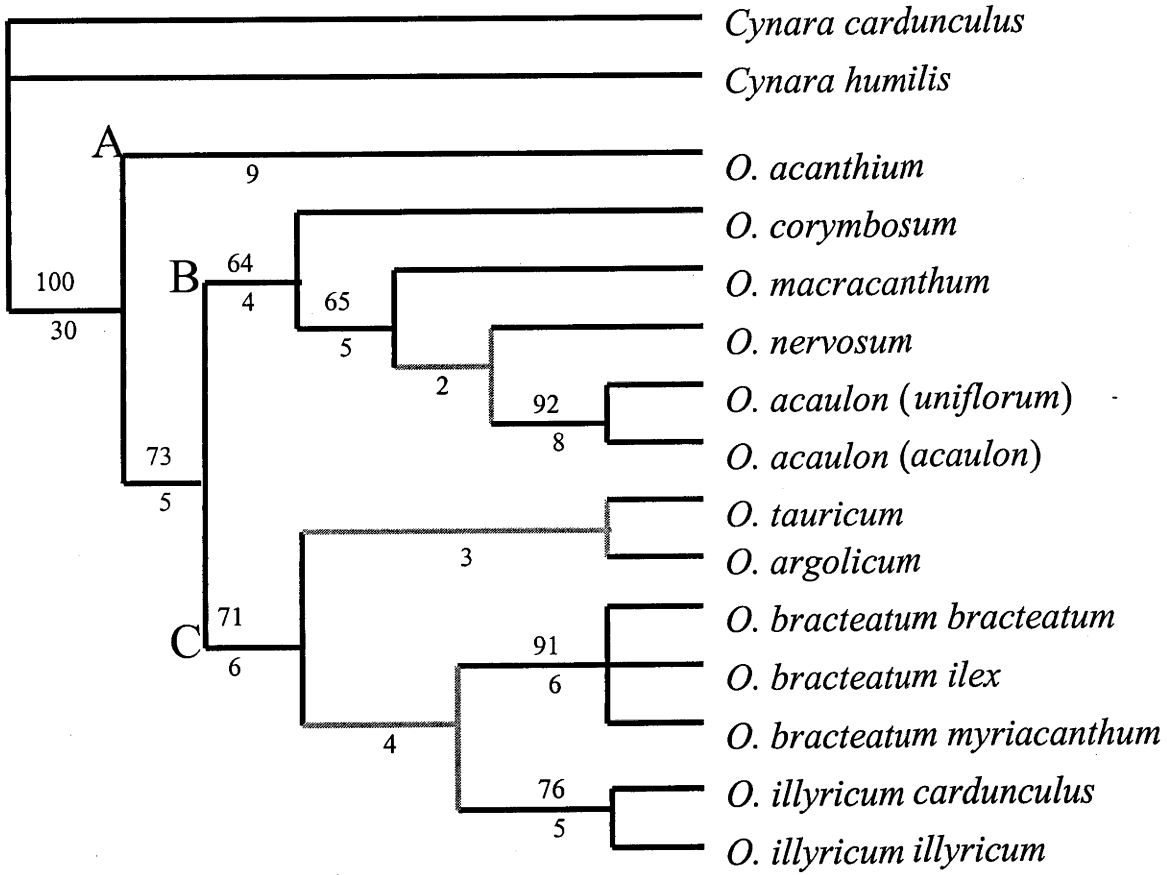


Figure 5.2 Split decomposition graph showing non-hierarchical relationships in the data. Groupings of taxa can be visualised by removing pairs of parallel lines. Samples that have more than one particular grouping are connected to other samples by more than a single vertex/join. The strength of particular groupings is illustrated by line length, showing the relative distance between samples, and bootstrap values presented on vertices (e.g. *O. illyricum* X *acanthium* is connected to *O. illyricum* as well as *O. acanthium*. The removal of one pair of parallel vertices with bootstrap values of 24% groups the sample with *O. illyricum* and the removal of one pair of parallel vertices with bootstrap values of 33% groups the sample with *O. acanthium* and other species).

Branch and bound parsimony analysis identified a single shortest tree of length 198 (HI 0.4365, RI 0.8277, RC 0.4682, Fig. 5.3). Bootstrap re-sampling of the data set produced a bootstrap 50% majority-rule consensus tree of length 215. Most branches were strongly supported by bootstrapping (Fig. 5.3). A neighbor-joining, bootstrap, 50% majority-rule consensus tree was computed using total character differences between taxa. The tree topology was identical to that obtained with parsimony analysis, and bootstrap values were also similar, so results are not shown here.



Morphological characters

	1	2	3	4	5	6	7	8	9	10	11
O. acanthium		5	35x40	0,6-8	7-9	2-4	14-25	a	s	u	2
O. corymbosum		5	40x10	6-8	8-10	3-4	20-25	a	s	m	1
O. macracanthum		6-7	40x20	5-7	7-9	5-6	25-30	a-r	s	u	2
O. nervosum		4	50x20	6-8	8-10	4-6	32-35	a	p	m	1
O. acaulon (uniflorum)		6-12	40x12	0	20-30	3-4	22-25	a	s	u	2
O. acaulon (acaulon)		6-12	40x12	0	20-30	3-4	22-25	a	s	u	2
O. tauricum		4	25x10	6-8	8-10	4-7	25-30	r	s	m	1
O. argolicum		5	25x6	6-8	8-10	5-7	25-30	a-r	s	m	1
O. bracteatum bracteatum		5	30x8	10-12	8-10	5-8	30-40	r	s	u	2
O. bracteatum ilex		12	30x8	10-12	8-10	5-8	30-40	r	s	u	2
O. bracteatum myriacanthum		5	30x8	10-12	8-10	5-8	30-40	r	s	u	2
O. illyricum cardunculus		2	55x15	8-10	10-12	5-7	30-35	r	p	u	2
O. illyricum illyricum		3	55x15	8-10	10-12	5-7	30-35	r	p	u	2

Figure 5.3 Strict consensus tree of European *Onopordum* species and distribution of selected morphological characters (Petit 1997): Bootstrap values given above branches. Grey branches represent bootstrap values <50%. A - Northern species (T-PTP non monophyly <0.01), B - Western species (T-PTP non monophyly <0.01), C - Eastern species (T-PTP non monophyly <0.01). Distribution of morphological characters: 1, Sections (light gray, Subgen. Acaulon; rest Subgen Onopordum: diagonal strips, ONOPORDUM; Dark gray, ECHINATA; horizontal stripes, ERECTA; and white, RECURVATA); 2, bract spine length (mm); 3, leaf maximum size cm^2; 4, pairs of leaf lobes; 5, pappus length (mm); 6, bract width (mm); 7, corolla length (mm); 8, bracts acuminate (a) or recurved (r); 9, pappus hairs scabrid (s) or plumose (p); 10, hairs on plant unicellular (u) or mulicellular (m); 11, Stem and mature leaves not tomentose or lanate (1) or densely white-grey tomentose or lanate (2).

Each species was found to be monophyletic. Within the tree several clades were revealed that conformed to the morphological classification of Franco (1976). The section RECURVATA here represented by *O. illyricum* and *O. bracteatum* was found to be natural on the basis of monophyly. In addition ONOPORDUM, here represented by *O. acanthium*, the only temperate species, also formed a unique clade and separated first from the remaining *Onopordum*. However, other clades comprised species not expected to be closely related. The species comprising clade B (*O. corymbosum*, *O. macracanthum*, *O. nervosum* and *O. acaulon*) all belong to different morphological sections in *Onopordum* (Franco, 1976). Morphologically, the two species, *O. acaulon* and *O. nervosum* are extremely different (*O. acaulon*; stemless, tomentose: *O. nervosum*; up to 3m tall, glabrous) but here they are revealed as the most closely related to each other, though the clade is supported by a bootstrap value of <50%. Furthermore, the section ECHINATA contained species present in both clades B and C, both of which are supported by moderate bootstrap values and T-PTP tests of monophyly (Fig. 5.3).

Discussion

Biogeographic analysis

Two centres of species diversity are present within Europe; around the Iberian Peninsula, and around the Aegean region (Franco 1976). *O. acanthium* is the most broadly distributed of *Onopordum* taxa, with *O. illyricum* also being present in much of the range around the Mediterranean basin. *O. corymbosum* is the most curiously distributed of the described morphological species as it has two disjunct areas of abundance, one in Northern Spain and one in the Balkans. However, most taxa are

restricted to one of the two major regions, suggesting that speciation has occurred within these areas. Phylogenetic reconstruction shows that there is a strong element of geographic structure to the relationships between species (Fig. 5.3) with clades A, B and C corresponding to different areas in Europe. Geographic proximity of species distributions provides a better estimation of phylogenetic relationships than morphological patterns. In both phylogenetic analysis and non-metric multidimensional scaling, a separation between species occurring in the Iberian Peninsula and those near the Aegean was obtained. Furthermore, only one species, *O. acanthium* occurs in the central and northern parts of Europe, and it was distinct from species found in these two groups (Figs. 5.1, 5.3).

cpDNA and AFLP polymorphism

Within *Onopordum* 127 polymorphic AFLP fragments were produced using a single primer pair combination, 25% of which were unique to individual species revealing both the rapid evolution of AFLP fragments, and the substantial degree of AFLP divergence between *Onopordum* species. AFLP variation therefore appears to be an appropriate marker for reconstructing relationships between *Onopordum* species. In contrast to AFLPs, no chloroplast variation was observed. It is difficult to infer from a single, but invariable, 800bp region of the chloroplast that there is no chloroplast divergence between the species studied. Other regions of the chloroplast may display more variation than *trnL-trnF* for this group and further work is required. However, as *trn*L-*trn*F has been shown to be a highly variable chloroplast region (Mes and Thart 1994; Kita *et al.* 1995; Gielly and Taberlet 1996; Kim *et al.* 1996; Fujii *et al.* 1997; Maguire *et al.* 1997; Mes *et al.* 1997; Sang *et al.* 1997; Cros *et al.* 1998;

Lanner 1998), some variation was expected. While *trn*L-*trn*F may be very slowly evolving in *Onopordum*, the lack of variation may also reflect a long history of hybridization and shared ancestry.

Phylogeny of Onopordum

Several novel findings have been observed in this study. Firstly I found that *O. acanthium* divereged before the remaining taxa and formed a clade consisting of the single species. The remaining species formed a strongly divereged group (5 character state changes, 73% bootstrap, T-PTP non-monophyly <0.001), which itself comprised two clades. Clade (B) is formed by *O. acaulon, O. nervosum, O. corymbosum*, and *O. macracanthum.* These species not only belong to different sections within *Onopordum* subgenus *Onopordum* (*O. nervosum* to ERECTA and *O. macracanthum* and *O. corymbosum to* ECHINATA) but to different subgenera, with *O. acaulon* being part of the subgenus *Acaulon*. This clade is supported by relatively high bootstrap value (78%) and T-PTP value of non-monophyly <0.001. Finally a third clade (C) consists of the subspecies of *O. bracteatum* and the subspecies of *O. illyricum* RECURVATA, as well as *O. argolicum* and *O. tauricum*, members of section ECHINATA. This too is supported by a relatively high bootstrap value (84%) and T-PTP value of non-monophyly of less than 0.001.

This phylogeny contrasts sharply with the one implied by the taxonomic classifications of Franco (1976) and Gonzalez Sierra *et al.* (1992). The subgenera *Onopordum* and *Acaulon* are the first division expected within the group. However, some species within subgenus *Onopordum* are more closely related to subgenus *Acaulon* than to other members of subgenus *Onopordum*. Several character states unique to *Acaulon* may therefore be derived. One characteristic, stemlessness, is a

relatively plastic character. In certain conditions this state can be induced in *O. acanthium* and perhaps also other species (pers. obs). The section ERECTA, here represented by a single species, is most closely related to subgenus *Acaulon*, though it is not possible to determine whether these two sections form a distinct clade without additional sampling of other species in the section ERECTA. Section RECURVATA, here represented by *O. illyricum* and *O. bracteatum*, is shown to be monophyletic. However, to make section ECHINATA monophyletic is unparsimonious. This section has been characterized by the presence of multi-cellular or unicellular hairs, the underside of leaves not being reticulately veined, glandular corolla-lobes, scabrid pappus-bristles, and involucral bracts being cochlearform or ovate-lanceolate narrowing into long, rigid pungent processes which are shorter than the florets (Franco 1976). Each of these characteristics separately exists in other sections of *Onopordum* but the combination of these characters has been used to define section ECHINATA (Franco 1976). Nevertheless, based on the AFLP result, both clades B and C contain members of ECHINATA and it appears therefore an unnatural grouping.

Widespread hybridization

Hybridization has long been invoked as an explanation for the presence of intermediates in *Onopordum* (Gonzalez Sierra *et al.* 1992). Hybridization has been confirmed by molecular techniques between *O. acanthium* and *O. illyricum* in Australia (*Chapter Five*, O'Hanlon *et al.* 1999). In the present study several plants of putative hybrid origin including *O. illyricum* X *O. acanthium* from Eyguirres, France, and *O. nervosum* X *O. acanthium* from Cuenca, Spain, were confirmed to be derived from inter-species mixing. In addition, samples from one population

suspected as representing 'pure' *O. bracteatum.* Specimens of *Onopordum bracteatum* obtained from Kastoria, Greece, were found to be derived from crosses between *O. bracteatum* and *O. acanthium.* While many of the species studied here generally exist in allopatry, contact zones occasionally exist between most, making widespread hybridization probable. Such hybridization is also likely to have occurred throughout the history of the genus, especially during periods of environmental change, when different species may have come into contact. Such a situation probably occurred during glacial maxima when their distribution was very restricted. As a contemporary analogue, in Australia *O. acanthium* and *O. illyricum* hybridize frequently due to the greater opportunities presented by the colonization of a new environment and the new ecological context in which the thistles occur (O'Hanlon *et al.* 1999).

It is also possible that widespread hybridization within the genus may have resulted in low levels of cpDNA divergence. As refugia may have been geographically restricted, the opportunity for chloroplast sharing and selective sweeps may have eliminated much of the chloroplast variation. Comes and Abbott (1999) found a similar discrepency between nuclear and chloroplast data in hybridizing *Senecio* and attributed that pattern to hybridization. However, only one chloroplast locus has been examined here and the particular non-coding regions within it may have evolved too slowly for variation to be observed. A hybrid origin for taxa within *Onopordum* is consistent with data from the guild of phytophagous insects that have specialised on this genus. While these insects were restricted to the genus, Briese *et al.* (1994) found little evidence for further ecological specialization at sub-generic levels, such as occurs in phytophagous insect faunas on other thistle genera.

Phylogeography

Clades A, B and C are each shown to be robust and to correspond closely with geographic distributions of species. Clade A contains the only species that exists in temperate Europe, clade B contains those species found in the Iberian Peninsula, and clade C corresponds to species occurring in the Aegean region. One morpho-species sampled from Spain (*O. corymbosum*) is also reported to occur in the Balkan region (Franco 1976). Analysis of samples from this latter region would enable testing of whether this population is really more closely related to Spanish species or is aligned with the other Aegean species. Similarly the presence of isolated populations of *O. illyricum* (an Aegean species by molecular determination) in Spain could be due to relatively recent range expansion, possibly human-assisted. For instance, populations of *O. tauricum* (an Aegean species) are known to have been introduced into southern France from Greece during the last century, where the species has become naturalised. Given the significant human and trade movements in the Mediterranean over the past 4000 years, it would be surprising if there were not recent range expansions of this type. More intensive sampling of populations of *O. illyricum* could help clarify these possibilities.

A divide between Aegean and Iberian taxa, whether at the subspecific or species level has previously been identified in many groups of Mediterranean plants (see Thompson 1999). That a similar divide has been revealed here probably reflects equivalent historical processes within *Onopordum*, and that the diversity present within the Iberian Peninsula and the Aegean region were once separated by glacial maxima. Such glacial maxima extended far south, and appropriate conditions for Mediterranean plants may have been found only at the very southern tips of the

continent, bounded to the south by the Mediterranean Sea. That the only species to exist in temperate Europe also forms a unique clade is also of interest. This species differs from the rest in that it favours temperate conditions and can be found in Northern Europe. As it separated before the remaining species, it may have either been derived out of the southern refugia, or more likely from separate more northern refugia, or from separate recolonization from Central Asia, where several other *Onopordum* species occur. Additional sampling of species from these regions would be required to determine which of these is most probable.

Morphological evidence could also support an Aegean / Iberian divide, but would argue that there were distinct ancestors of species groups, such as *O. tauricum* / *O. corymbosum* and *O. argolicum* / *O. macracanthum*, that became separated during retreat into glacial refugia and diverged. The novelty of the molecular data set is that it indicates that isolation of the Aegean and Iberian taxa preceded the most recent major speciation in this group in Europe. Again, support for this comes from a study of the phytophagous insect fauna on *Onopordum*. Briese *et al.* (1994) found that the composition of the specialist fauna is likely to be more similar between species from the same region than between the above groups of apparently related but isolated species. Also, speciation within a group of weevils of the genus *Larinus*, specialist endophages on *Onopordum* capitula has been shown to have been driven by biogeographic isolation between the Iberian and Aegean regions rather than host-plant specialization within the genus (Briese *et al.* 1996).

Convergent morphology and hybrid refugia

As shown previously, clade ECHINATA has representatives in both Eastern and Western distributions. However, they do not share a most recent common ancestor

and representatives from each region are more closely related to other species from that region. This presents an interesting case of apparent convergent evolution. As many morphological characters are uncorrelated with the phylogeny (as well as with morphological classification) (Fig. 5.3), it is unlikely that they all evolved separately on several occasions. Rather, it is more probable that this range of traits were present prior to speciation. Such a pattern could emerge through a combination of vicariance events of glaciation, hybridization, and subsequent sorting of these traits, especially since *Onopordum* species are prone to hybridize. The fact that the morphologically similar, yet genetically different *O. argolicum* and *O. macracanthum* are both restricted to more arid extremes of their respective Aegean and Iberian ranges supports this idea. Hybridization has also been invoked as the mechanism explaining a similar conflict between morphological and molecular phylogenies in the genus *Armeria* (Plumbaginaceae) in the Iberian peninsula (Fuertes Alguilar *et al.* 1999).

Given the contemporary distributions of most *Onopordum* species, it is evident that tundra would have eliminated any species incapable of persisting in the southern refugia. As those refugia would have been small, and consequently the probability for inter-specific gene flow within them would have been high, refugia may have been abundant with hybridizing populations, and hence contain high proportion of the genetic diversity present before glaciation. After glacial retreat and a northern advancement by hybrids, the opportunity for the sorting of traits by selection could be producing the pattern observed today. The complexity of this pattern and the weak internal reproductive barriers that exist within *Onopordum* make conventional taxonomic treatment of this genus a difficult task. It is clear from these data that such systematic studies should examine phylogenetic relationships at morphological and molecular levels as part of any attempt to delineate taxa.

Chapter Six *AFLP reveals introgression in weedy* Onopordum *thistles - hybridization and invasion*[3]

***Onopordum* L. (*Asteraceae*) is an extremely diverse genus of thistles that includes several species that have become serious pasture weeds in several regions of the world. I present a comparison of the genetic diversity in invasive forms of *Onopordum* from Australia with several known native European species. A total of 108 polymorphic genetic markers were generated using AFLP fingerprinting. Non-metric multidimensional scaling (NMDS) revealed that Australia contained *O. acanthium*, *O. illyricum* and a full range of genetic intermediates between these species. Intermediate genotypes were largely made up of segregating fragments diagnostic for European *O. acanthium* and *O. illyricum* with a low frequency of fragments that were diagnostic for other species never recorded in Australia. The current genetic patterns in Australia may be best explained by a combination of processes, both in the native and in the alien range. These include multiple introduction of seed, including hybrid material, and the continuous dispersal in Australia, leading to an increase in the contact among hybridizing taxa. Such processes appear to have produced more widespread hybridization and introgression in Australian *Onopordum* than is found in Europe.**

It is now established that hybridization is both a widespread and potentially creative evolutionary phenomenon (Rieseberg 1997). The outcome of hybridization for speciation depends on the fitness of hybrids (Emms & Arnold 1997), which can vary with environmental conditions. However, hybrids are not uniformly fit (Arnold & Hodges 1995), and environmental conditions are dynamic. Invading species provide an interesting setting in which species capable of hybridizing may come together in a disturbed or novel context. An important case of the synergy between the invasion of a novel environment and hybridization is found in *Carduus* thistles. Two allopatric species native to Europe (*C. nutans* and *C. acanthoides*) hybridized when introduced together into Canada, resulting in the local abundance of hybrid and backcrossed plants (Warwick *et al.* 1989).

[3] This chapter is an expansion on O'Hanlon, Peakall & Briese (1999)

Species of another thistle genus, *Onopordum* (Asteraceae) have been introduced into Australia, North America, South America and New Zealand, where they have become major pasture weeds. *Onopordum* is of Eurasian origin with over 100 species described (Dress 1966). However, taxonomic re-evaluation suggests that many described species are synonymous and the actual number is probably closer to 50 (Michael 1996). In Europe, about twenty species occur, mostly in allopatry (Franco 1976, Briese 1990). Only rarely do species come into contact, and then only at low population densities. At such locations, morphological analysis has revealed plants of intermediate morphology, indicating that hybridization between species is possible (Gonzalez Sierra *et al.* 1992).

Onopordum species appear to have been introduced into Australia on many occasions, both as an ornamental plant and as an agricultural contaminant (Briese 1990). However, only four described species have become naturalized; *Onopordum acanthium* L., *Onopordum illyricum* L., *Onopordum acaulon* L. and *Onopordum tauricum* Willd. (Briese *et al.* 1990), all of European origin. *O. tauricum* has been observed only a few times in Australia, and then, only in localized situations. In contrast, *O. acaulon* occupies much of southern Australia. It is the most widespread of the species, and occurs in the absence of other *Onopordum* species. It is not, however, considered as serious a weed as either *O. acanthium* and *O. illyricum* which form very dense infestations. Together, these two species infest over 1,100,000 hectares, and have been the subject of co-ordinated control for several years (Briese 1990). They are patchily distributed, consistent with their lack of wind dispersal, and often occur in dense clumps. While several infestations contain only *O. acanthium* or *O. illyricum*, field observations indicate that they sometimes co-

occur. Furthermore, the presence of additional taxonomically ambiguous but morphologically intermediate populations have been identified. These intermediates also occur in the absence of other *Onopordum* species over large parts of the distribution (Briese *et al.* 1990, Groves *et al.* 1990). Samples from such populations do not correspond with type specimens of any *Onopordum* species (Michael 1996). Because of their intermediate nature, and the suggestion of hybridization between *Onopordum* species in native habitats, it has been proposed that the intermediates may have arisen through matings between *O. acanthium* and *O. illyricum*, leading to hybridization or introgression (Michael 1968, Groves *et al.* 1990, Parsons & Cuthbertson 1992). Given the lack of information about these populations, it is also possible that they are the product of matings between several species. Such multi-species introgression between European *Onopordum* may have occurred prior to invasion and establishment as a weed in Australia.

Apart from illuminating evolutionary processes, there is a potential practical aspect to the study of the evolution of invading weeds. Effective weed management requires an understanding of weed diversity, at the genetic as well as at the physiological and ecological levels (Nissen *et al.* 1995, Jasieniuk *et al.* 1996, Jordan & Jannink 1997). Hence for the most effective long-term management of *Onopordum*, an understanding of taxonomic relationships, their evolutionary history, and consequent spread in Australia is required (Briese *et al.* 1990, Groves *et al.* 1990, Michael 1996). Furthermore, if hypotheses about the hybrid origin of *Onopordum* populations are confirmed, this system may serve as an example of the importance of both hybridization and environmental variation in the evolution of alien taxa.

In this chapter I describe the genetic diversity in Australian forms of weedy

Onopordum by comparison with reference material of European *Onopordum*. I show that plants of intermediate morphology between *O. acanthium* and *O. illyricum* are also genetically intermediate, with the pattern of intermediacy being consistent with hybridization and inconsistent with either convergence or plesiomorphy. I then explore the evidence for multi-species introgression and assess whether hybridization occurred before or after introduction to Australia.

Materials and Methods

Sampling

To serve as reference material, several species of European *Onopordum,* including *O. acanthium* and *O.illyricum,* were sampled for genetic analysis (Table 6.1). These species were identified by an independent taxonomic expert, J. Amaral do Franco, based on careful comparison with the descriptions and type voucher specimens described in Franco (1976). Voucher samples of seed from the samples collected are kept at CSIRO Entomology, Canberra, Australia. These twenty-five samples served as a base-line with which to describe the genetic diversity in Australian samples of *Onopordum,* with all European species except for the highly localized *O. majorii* Beauverd., *O. messianicum* Halacsy., *O. laconicum* Heldr. and Sart., and *O. caulescens* D'Urv. (Franco, 1976) being included.

Table 6.1 Species and sites sampled as reference material in *Onopordum* native range.

Species	Subspecies	Location
O. acaulon L.	*acaulon*	Sierra Segura, Spain
		Zaragoza, Spain
	uniflorum (Cav.)	Llo-E. Pyrenees, France
O. acanthium L.		Cuenca, Spain
		Florina, Greece
		Tobia, Italy
		Eyguieres, France
O. macracanthum Schousboe		Cordoba, Spain
O.argolicum Boiss.		Tripoli, Greece
O. corymbosum Willk.		Teruel, Spain
O. tauricum Willd.		Lake Volvi, Greece
O. bracteatum Boiss. & Heldr. In Boiss.	*bracteatum*	Kastoria, Greece
	ilex (Janka) Franco	Lamia, Greece
	myriacanthum (Boiss.) Franco	Argos, Greece
		Epidavros, Greece
O. illyricum L.	*Illyricum*	Tobia, Italy
	illyricum	Bonifacio, Corsica
	cardunculus (Boiss.) Franco	Igoumenitsa, Greece
		Ionnina, Greece
O. nervosum Boiss.		Cuenca, Spain

The level of interest here is in taxonomic, as opposed to population differences. Plant specimens and seeds were collected during the summer of 1992 from 24 sites ranging from northern New South Wales to central Victoria. In 1996, seedlings were grown from 24 of these sites (Table 6.2, Fig. 6.1) representing the widest geographic distribution and the full range of morphological variation (Table 6.2). Two seedlings, derived from different parents were harvested from each site for genetic analysis and a single pressed specimen was used for morphological analysis. While these values are small, the goal was not to partition variation within and among sites, but rather, to describe the overall pattern of diversity present in Australia. While a single sample from each site would have been adequate for my purposes, I have analyzed an additional sample from each site in an attempt to recover a greater degree of the overall genetic diversity. Seeds from each site were germinated to give a total of 48

individuals, with two seedlings, derived from different parents being harvested from each site.

Morphological analysis

Several characters including leaf size and shape, stem architecture, the number and size of flowering heads and bract size and shape were highly variable and were used to classify field samples into *a-priori* phenotypes (Table 6.2). The purpose of defining such groups was to aid graphical presentation rather than to taxonomically define samples. The groups included *O. illyricum*, *O. acanthium*, and putative introgressants.

Variation in the spine-tipped bracts surrounding flower heads is genetically determined and used widely for taxonomic purposes in both the tribe and the genus (Franco 1976, Petit 1997). Bracts are particularly useful as they are readily preserved in vouchers and are easily measured, giving adequate information for cross-species comparisons. Consequently, previous collections are biased toward the sampling and preservation of bracts. Bracts are variable in many features, and three measurements were used for this investigation; maximum width, length and shape. These characters readily separate *O. acanthium* and *O .illyricum*. In order to avoid developmental variation, bracts from the third outer seires were measured to the nearest half millimetre. Shape was determined as the width at the base divided by the width at the mid-point divided by the width at two-thirds of the length for each bract. This enabled the determination of bract shape independent of size. *O. acanthium* quickly tapers from the base and then gradually narrows to the tip, whereas *O. illyricum* remains wide from the base and narrows only near the tip. Putative hybrids taper

gradually from the base to the tip. To summarise the morphological variation, a Gower distance metric was employed to measure the pairwise distance between individuals from each site using SYN-TAX (Podani 1995).

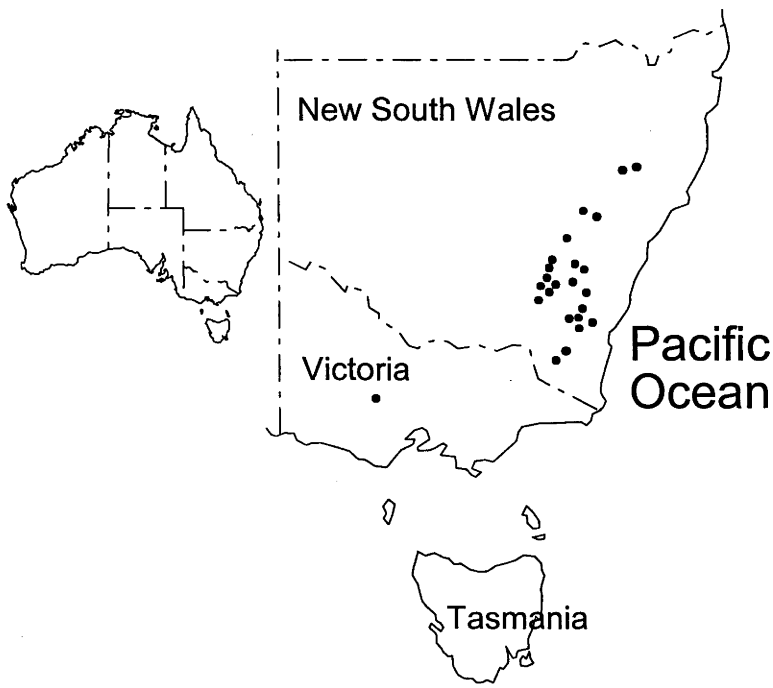


Figure 6.1 Twenty-four sites sampled in southeastern Asutralia for amplified fragment length polymorphism (AFLP) analysis.

Genetic analysis

Amplified Fragment Length Polymorphism (AFLP) is a relatively new procedure (Lin & Kuo 1995, Vos *et al.* 1995) that has already been applied widely to a range of agricultural and native plant species. Recent studies show that AFLP is a reliable method for generating many polymorphic markers (Powell *et al.* 1996, Maheswaran *et al.* 1997, Krauss & Peakall 1998). Total genomic DNA was extracted and purified from seedlings which were ground in liquid nitrogen and purified of protein, RNA and polysaccharide using the diatomite method of Gilmore *et al.* (1992). DNA was precipitated in ethanol and resuspended in 50uL of low TE (10mM Tris-HCl (pH 8.0), 0.1 mM EDTA) buffer. AFLP fingerprints were generated according to the procedures contained in *Chapter Three.*

Statistical analysis of genetic data

The relative mobility of fragments was accurately calculated by the inclusion of an internal size standard within each sample. Digital profiles were visualized and scored with the aid of ABI GeneScan software. Fragment differences between samples were identified as those fragments differing by at least 0.5 relative migration units (bp) in two separate reactions. Binary vectors for the presence or absence of fragments were constructed for each sample and used to calculate a pairwise genetic distance matrix using the Squared Euclidean distance metric (Huff *et al.* 1992) in the software package SYN-TAX (Podani 1995).

The pairwise genetic distance matrix was then analyzed using two dimensional non-metric multidimensional scaling (NMDS) in SYN-TAX (Podani 1995). NMDS is an ordination procedure that finds those positions of *n* individuals in reduced space that best reflects their original genetic distances. It is an iterative process using a steepest descent algorithm to minimize stress, the deviation between the final distance matrix to the original one (Kruskal 1964). It is useful for recovering non-hierarchical structure while avoiding assumptions of the identity of individuals, normal data, or the underlying pattern of structure within the data. NMDS also provides a better fit to the data than may be obtained using other ordination techniques (Lessa 1990, Krauss 1996). NMDS was performed 30 times from an initial random configuration of individuals in two-dimensional space, to minimize the possibility of finding a local rather than a global minimum. A further iteration was performed using the resemblance matrix from principal coordinated analysis (PCO) as an intital configuration for iteration to test the relative fit of PCO and NMDS to the original

distance matrix.

Species contributions in Australian Onopordum

High generation hybrids and later generation backcrosses may contain only a small number of fragments from one or more of the contributing species. NMDS summarises the overall differences between samples but cannot reveal which species contribute to the DNA profile of a particular sample. Therefore, the frequency of individual AFLP fragments present in Australian *Onopordum* that were diagnostic for different European species were determined. To do this, AFLP fragments were assigned to one of five categories: 1. fragments diagnostic for *O. acanthium*, 2. fragments diagnostic for *O. illyricum*, 3. fragments present in both *O. acanthium* and *O. illyricum* but absent from all other species, 4. fragments diagnostic of species other than *O. acanthium* or *O. illyricum*, 5. fragments present only in Australian *Onopordum* and therefore absent from the European *Onopordum* samples. In the final analysis, the average percent frequency of the different categories of AFLP fragments was determined for three groups of Australian *Onopordum*; *O. acanthium* defined by the presence of only *O. acanthium* fragments, *O. illyricum* defined by the presence of only *O. illyricum* fragments and hybrids defined by the presence of both *O. acanthium* and *O. illyricum* fragments.

Morphological genetic correlation

Floate *et al.* (1994) indicate that many markers are required to adequately classify hybrid plants, whether morphological or genetic characters are used. Furthermore, a combination of both types of markers may be the most useful. In order to test for correlation between data sets, and to assess the predictive value of the much simpler

morphological data for genetic identity, matrix correlation between genetic and morphological distance matrices, and transformed rank genetic and morphological distance matrices were performed using SYN-TAX (Podani 1995). Since the genetic distance matrix contained twice the number of individuals as the morphological distance matrix, it was reduced in size by taking the average pairwise genetic distance between all sites.

Results

Morphological analysis

Pressed samples of *O. illyricum* and *O. acanthium* could easily be distinguished using the three morphological characters devised from the bract variation (Table 6.2). Samples from intermediate sites were mostly intermediate between populations putatively identified as containing parental species, with degree of overlap in most characters. With the use of only three characters of unknown inheritance, it was not possible to determine whether this plot reflected hybridization, introgressive hybridization or simply morphological intermediacy.

Table 6.2 Phenotypes presented; A (*O. acanthium*), I (*O. illyricum*) and X (putative hybrid) were determined from gross observation. Bract width defined as the maximum width (mm), bract length is defined as length from base to tip (mm) and bract shape according to formula in text.

Site (nearest town)	Latitude	Longitude	Phenotype	Bract width	Bract length	Bract shape
Mudgee	32°36'	149°35'	A	3.5	21.0	3.5
Lue	32°40'	149°51'	A	3.0	17.0	3.0
Nundle	31°27'	151°08'	A	2.5	14.5	2.5
Berridale	36°22'	148°50'	A	4.0	16.0	2.0
Cooma	36°14'	149°08'	A	3.5	16.0	2.0
Orange	33°17'	149o06'	A	3.5	18.5	1.8
Cowra	33°50'	148°41'	A	2.5	16.0	1.3
Koorawatha	34°02'	148°34'	A	2.5	16.0	1.3
Pialligo	35°17'	149°13'	A	3.5	16.5	0.8
Braidwood	35°27'	149°48'	X	4.5	26.0	2.3
Boorowa	34°27'	148°44'	X	4.0	21.5	1.5
Murringo	34°18'	148°31'	X	3.5	21.0	1.4
Kingsdale	34°40'	149°40'	X	2.5	16.0	1.3
Captains Flat	35°35'	149°27'	X	5.0	25.0	1.1
Bungendore	35°15'	149°27'	X	4.5	25.0	0.9
Binda	34°20'	149°22'	X	3.5	17.5	0.8
Tuena	34°01'	149°20'	X	8.0	32.5	0.7
Murrumbarra	34°33'	148°21'	X	6.0	28.0	0.7
Jugiong	34°50'	148°19'	X	6.0	25.0	0.7
Peelwood	34°07'	149°26'	X	7.5	32.5	0.5
Galong	34°36'	148°34'	X	7.5	31.5	0.5
Nundle	31°27'	151°08'	I	7.5	29.0	0.7
Quirindi	31°30'	150°41'	I	8.0	33.5	0.5
Maryborough	37°03'	143°44'	I	9.0	36.0	0.2

AFLP polymorphism and genetic analysis

Each sample generated 29.2 (s.e.= 1.5) fragments within the size range of 75-450bp. With a single AFLP primer pair combination, a total of 143 fragments were identified, 108 of which were polymorphic. NMDS of the genetic distance matrix produced a plot with a stress value of 13.2% deviation between the final NMDS result, and the original distance matrix. Such a value compared favourably with that produced by PCO, which showed a poorer fit to the data (Stress = 24.3%). Figure 6.2 shows that all species defined by European samples were easily distinguished from one-another. By contrast, samples collected from Australia fell along an axis separating *O. acanthium* and *O. illyricum* . Notably, the Australian samples spanned the full range between these species (Fig. 6.2).

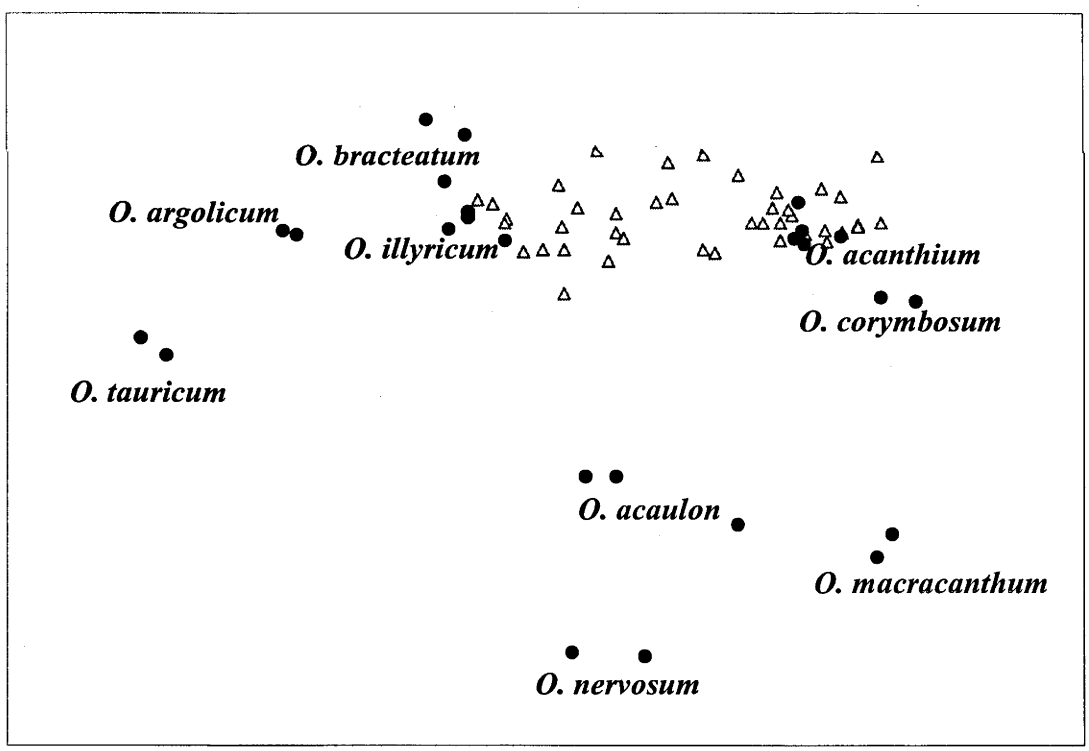


Figure 6.2 Two dimensional NMDS scatterplot of pairwise genetic distances between all 80 samples. Open triangles represent individuals sampled from their alien range (Australia) and closed circles represent individuals of known identity, sampled from their native European range.

Fragment frequencies

Three Australian locations were found to contain individuals with AFLP fragments diagnostic only of *O. acanthium*. The absence of any diagnostic fragments for other *Onopordum* species suggest that these populations are true *O. acanthium*. However only one location in Australia contained samples with AFLP fragments characteristic of European *O. illyricum*. The remaining forty samples, from twenty sites, all contained a combination of fragments diagnostic for both *O. acanthium* and *O. illyricum*, as well as some fragments diagnostic of several other species not found in Australia. The putative hybrid populations also contained a low proportion of fragments that could not be traced to any European samples of *Onopordum*. While it is possible that such fragments are the result of recombination in hybrids and large scale post-hybridization genetic re-organization as apparent in other hybrid species such as *Helianthus* L. sunflowers (Rieseberg, 1997), it is more likely that the presence of such fragments reflects the limited sampling within European species.

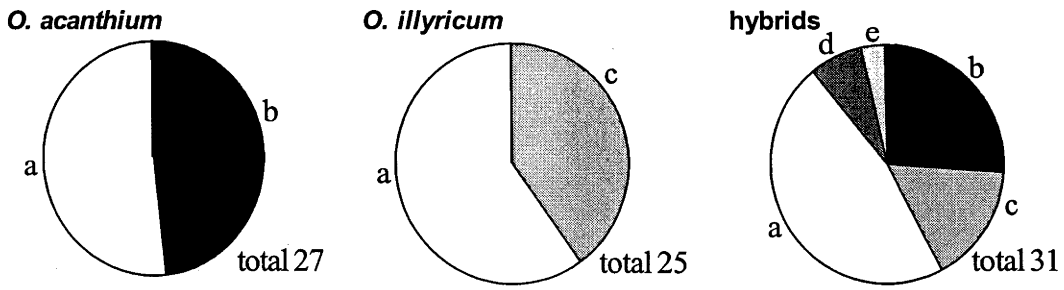


Figure 6.3 Contribution of AFLP fragments to Australian populations from European reference samples (a, present in both *O. acanthium* and *O. illyricum*, b, present in only *O. acanthium*, c, present in only *O. illyricum*, d, present only in species other than *O. acanthium* and *O. illyricum*, e, not present in European samples). The average total number of fragments per individual for each of the three groups is also given. Number of sites (*O. acanthium* 3, *O. illyricum* 1, *hybrids* 20).
*Australian populations of *O. acanthium* were identified by the absence of fragments diagnostic for any other species, and similarly for *O. illyricum*.

Correlation between data sets

The morphological and genetic AFLP data were significantly correlated (m=0.457, $p<0.001$). This correlation is surprisingly high considering the relative simplicity of

the morphological data. The findings indicate that bract morphology alone was predictive of the degree of genetic intermediacy between *O. acanthium* and *O. illyricum* and these data show a high potential for refinement of morphological analysis. However, character choice will be important in obtaining data that correlates to the genetic data as more characters will not necessarily produce a higher correlation.

Discussion

Patterns of Diversity

This study has confirmed the presence of taxa in south-eastern Australia that correspond closely to European *O. acanthium* and *O. illyricum*. However, populations of these species are far outnumbered by populations of genetic intermediates. Furthermore, the relative genetic distances between the intermediates show that they form a very genetically diverse group, being distributed along an axis between *O. acanthium* and *O. illyricum* (Fig. 6.2). The high degree of genetic diversity within the hybrid groups is reflected by the contribution of fragments diagnostic for both *O. acanthium* and *O. illyricum* (Fig. 6.3). While the variation is substantial, a relatively accurate classification of plants in the field may be obtained by conducting simple morphological measurements of bract variation and comparison with the values given in Table 6.2.

Hybridization has often been invoked as an explanation for intermediacy, although often without adequate evidence (Rieseberg 1997). Morphological or genetic intermediacy alone may be a reflection of convergence, where a third species

independently acquires characters similar to those of putative parental species. It may also reflect plesiomorphy, or the retention of ancestral character states by an intermediate group, with differentiation giving rise to apparent parental taxa. (Rieseberg & Ellstrand 1993, Allan *et al.* 1997, Rieseberg 1997). It is crucial to determine which of these explanations best explains the observed patterns in my study of Australian *Onopordum*.

Hybridization, Convergence or Plesiomorphy

Evolutionary time is required for the patterns produced by convergence or plesiomorphy to emerge, and these processes are unlikely to have occurred within the 150 years since the introduction of *Onopordum* into Australia. They are also unlikely to have occurred prior to invasion. *O. illyricum* and *O. acanthium* are thought to have diverged relatively early, and belong to different sections within *Onopordum* (Franco 1976, confirmed *Chapter Five*) and no extant morphologically intermediate species are known. The data provided here also shows that in Europe, none of the described species are genetically intermediate between these species (Fig. 6.2). These species are both more closely related to several other species than to one another. Therefore, it is unlikely that the source of the intermediate group in Australia was an extant ancestral species, and plesiomorphy is therefore improbable. Pre-colonization convergence is also an unlikely explanation for the patterns of diversity in Australian *Onopordum*. Due to the large number of polymorphic nuclear fragments that segregate in putative hybrids, there is a very low probability that the intermediates could have acquired all of these fragments through the normal processes of convergent evolution. Since convergence and plesiomorphy are unlikely, hybridization most adequately explains the current patterns of genetic variation in

Australian *Onopordum*. However, hybridization may have occurred both before and after invasion.

The Timing of Hybridization

Three lines of evidence suggest that pre-invasion hybridization was possible. Firstly, putative natural hybrids between *O. acanthium* and *O. nervosum*, *O. acanthium* and *O. tauricum*, *O. acanthium* and *O. illyricum*, *O. acanthium* and *O. acaulon*, *O. illyricum* and *O. nervosum, O. nervosum* and *O. tauricum*, *O. nervosum* and *O. macracanthum* have been reported in the native range (Gonzalez Sierra, 1992). Secondly, at least some introduced seed is suspected to have been deliberately introduced from botanic gardens where different *Onopordum* species grow in cultivation together, and hybridization is likely to be frequent. Finally, several fragments diagnostic of *Onopordum* species not found in Australia were found in Australian intermediates (Fig. 6.3). These findings may best be explained by multi-species introgression, which can only have occurred prior to invasion.

Given that there is good evidence for hybridization in Europe, it is apparent that opportunities for hybridization in Australia may have been even greater. The current genetic patterns may be best explained by a combination of processes, both prior to and after invasion of Australia. These include multiple introductions of seed, including hybrid seed, and continuous dispersal which increased the contact among hybridizing taxa. Such processes appear to have produced more widespread hybridization and introgression in Australian *Onopordum* than is found in Europe.

Evolutionary Implications

There has been debate about the long-term importance of hybridization to evolution (see Anderson 1949, Harrison 1993). However, several well documented studies have provided evidence in support of its importance (Arnold & Bennett 1993, Rieseberg & Wendel 1993, Rieseberg 1997). Where geographic boundaries, rather than biological mechanisms of reproductive isolation have led to speciation, and when species colonize a new region, the scope for hybridization is high. Hybridization has previously been confirmed between other invading species such as *Senecio* (Abbott 1992) and *Carduus* (Warwick *et al.* 1989). These findings for *Onopordum* thistles provide an additional demonstration of extensive introgressive hybridization between invading weeds.

Since colonizing Australia, *Onopordum* has been faced with a suite of ecological conditions different than those in Europe. Apart from gross climatic, edaphic and pasture composition differences, European *Onopordum* serves as a host to many herbivorous insect species absent from Australian populations (Briese 1990). The relative abundance of hybrids in Australia may indicate that they are better adapted to the novel Australian conditions than they are to European ones. Conversely, the apparent fitness may indicate that hybrids have been released from conditions selecting against them, as predicted by the tension-zone (Barton & Hewitt 1985) and ecological gradient models (Moore & Price 1993). On the other hand, the simplest explanation for the abundance of *Onopordum* hybrids may simply be a consequence of their wider introduction and dispersal. Clearly, determining which of these hypotheses is correct will aid in understanding the importance of hybridization to the evolution of *Onopordum,* a genus that may provide an example of where

hybridization has aided adaptation to new environmental conditions.

Implications for Weed Management

Recent attention has focused on the consequences of plant hybridization to insect herbivores. Currently, there appears no general response by herbivores to plant hybrid zones with results ranging from decreased to increased herbivory (Floate *et al.* 1997, Fritz *et al.* 1997, Whitham *et al.* 1997). Introgressive hybridization could potentially provide "bridges" for insects to switch hosts (Marohasy 1996, Fritz *et al.* 1997, Holt & Hochberg 1997). However, to date there is little evidence of host-shifting. Similarly there has been suggestions that hybridization may lead to the spread of herbicide resistance in weeds (Jasieniuk *et al.* 1996).

While the genetic consequences of hybridization for host-specificity, herbicide control and other weed management strategies is unknown, there are several ecological and morphological factors distinguishing *O. acanthium* and *O. illyricum* which highlight the relevance of these possibilities. Variable characters such as leaf pubescence, facultative perennation, flowering predictability, flower head size and temperature affiliation, may influence the success of particular control strategies (O'Hanlon *et al.* 2000)[b]. Biological control (Briese 1990), pasture management, (Pettit *et al.* 1996) and herbicide application and effectiveness (Peirce 1996) may all be influenced by such variation. The probabilities that such variation exists in Australian *Onopordum* is increased by the fact that species other than the putative parents appear to be involved in multi-species introgression. Because the present study has shown that hybridization occurs in *Onopordum*, it is important to test whether all control techniques work on the full range of this diversity. To date

Chapter Seven *Colonizing founder effects in hybrid Onopordum thistles - molecular and morphological analyses*

It is commonly held that hybrid speciation is a special case of sympatric evolution as hybridization must occur in the presence of parental species. However, the isolation of hybrid lineages from one-another and from parental species may be possible, and perhaps common where hybrid founders invade new localities. Indeed, periods of rapid hybrid speciation through allopatric processes may accompany range expansion, often associated with environmental change. I report an example of widespread introgression in *Onopordum* spp. (Scotch thistle) infestations in south-east Australia. AFLPs were used to generate 29 species diagnostic markers for 98 individuals from 7 populations. Analysis of molecular variance and non-metric multidimensional scaling demonstrated that most infestations were of low genetic diversity, as well as differentiated from each other resulting in a suite of isolated hybrid lineages. Furthermore, spatial autocorrelation analysis of morphological data from 76 newly established populations suggested that hybrid *Onopordum* infestations were often established by secondary dispersal from neighboring sites rather than by primary contact between parental species. *Onopordum* therefore represents an example of the importance of colonisation to hybrid speciation, and taxa in which the requirements for early allopatric hybrid speciation have been met.

Speciation by homoploid hybridization has been considered theoretically difficult to achieve because of the problems associated with evolving reproductive isolation during ongoing genetic exchange with parental species (Barton & Hewitt 1985, Harrison 1993, Rieseberg & Carney 1998). However, the recombinational model of homoploid hybrid speciation is now widely accepted as plausible, and has been demonstrated in sunflowers (Rieseberg 1997). In brief, such a model predicts that in appropriate circumstances, the sorting of parental chromosomal and genic sterility factors will lead to the formation of a fertile hybrid derivative that is at least partially intersterile with its progenitors. Habitat dependent selection then favors the most viable of the hybrid segregants producing a relatively fit hybrid neospecies (Emms and Arnold 1997, Arnold and Hodges 1995 Rieseberg and Carney, 1998). However, in certain cases hybrid vigor alone may result in the stabilization of hybrid lineages

(Arnold & Bennett 1992, Arnold & Hodges 1995). Generally F_1 hybrids between locally adapted races or closely related species are more vigorous than their parental progenitors. If such hybrid vigor (heterosis) is maintained through successive generations, such a phenomenon could result in the replacement of parents by hybrids (Levin 1996, Rieseberg and Carney, 1998). For this to take place, strong selection for hybrids in a unique habitat is usually required (Anderson 1949, Emms & Arnold 1997, Rieseberg & Carney 1998), and hybrid vigor is rarely maintained in segregating generations (Anderson 1949, Arnold & Hodges 1995, Rieseberg & Carney 1998).

Hybrid speciation may be simpler to achieve in the absence of parental species, without ongoing gene flow (Charlesworth 1995, Rieseberg & Carney, 1998). If hybrids and parents are spatially separated, the criteria of either reproductive isolation or hybrid vigor need not be met, and the isolated hybrid lineage may evolve as any other normal population might. Colonization possibly provides more opportunities for hybrids to become isolated from parents (Abbott 1992, Charlesworth 1995), especially as self fertile hybrids require only one founding event at a given location to establish an independent lineage.

Despite this possibility, there are few examples of hybrids existing in isolation from parents. This may be because hybrids are often only identifiable with a disjunct hybrid zone (Charlesworth 1995, Rieseberg 1997). For hybrids between the thistles *Carduus nutans* and *C. acanthoides*, a population containing only introgressed individuals was found near a hybrid swarm, and retained its independent identity for a period of at least 30 years (Warwick *et al.* 1989). While it was not possible to determine whether this was due to selection or lack of migration, it is likely that it was founded by a

backcross from a neighboring population. A study of knotweed species (*Fallopia*) in the United Kingdom revealed that hybrid knotweed occasionally co-occurred with parental species, but more often in isolation, possibly through hybrid founders or hybrid vigor. For the establishment of hybrid neospecies in allopatry, hybrids must (i) not only be isolated from parents but (ii) the isolation must be brought about by hybrid founders, and not through fitness effects and displacement of parents. In this chapter, I demonstrate that both of these conditions have been met in weedy *Onopordum* thistles. This system is undoubtedly semi-natural, and much of the dispersal is human assisted. However, *Onopordum* thistles display occasional long distance animal-assisted dispersal in their native range and have become adapted to pastures grazed by large animals. Consequently, it is possible that such a pattern could emerge without human intervention, especially during episodic climatic changes such as glacial retreats where many plants are known to have migrated very rapidly as described by Hewitt (1999).

Previously, in a broad scale genetic study of *Onopordum* thistles, two samples each from 24 populations from south-east Australia were analyzed, showing that only two populations contained *O. illyricum* and three populations contained *O. acanthium*, with all remaining populations containing plants only of hybrid origin (O'Hanlon *et al.* 1999; see also *Chapter Six*). Similarly to studies in other taxa, it was not known whether these populations were the result of *in-situ* contact between parental species or whether they were derived from founding events by hybrids. Due to limited sampling within populations, it was also unclear whether these populations were distinct from each other or whether ongoing gene flow between populations meant that hybrids were not isolated from parents (O'Hanlon *et al.* 1999; *Chapter Six*). In

addition, it was not possible to determine whether the parental and hybrid species were randomly distributed, or structured through either dispersal patterns or by selection.

In this chapter, I confirm with morphological data that the distribution of these species in Australia is non-random. Secondly, *via* spatial autocorrelation analysis, I demonstrate that contemporary patterns of morphological variation have been determined to a large degree by local dispersal, and founding events from neighboring sites, as well as occasional long distance dispersal. Finally, as it has previously been shown that many sites may contain hybrids which can act as founders for other populations, and as previous small scale genetic studies of within population variance indicated that many hybrid sites contained no parental species (*Chapter Six*), I document a more intensive genetic study within selected populations confirming that a suite of hybrid lineages has become isolated from parents. Given these findings, *Onopordum* thistles appear to be an effective demonstration that the interaction between invasion and hybridization can produce the necessary precursors for the evolution of a suite of distinct lineages.

Materials and Methods

Morphological analysis

A large scale morphological analysis across the species' distributions in New South Wales, Australia (Fig. 7.1) was conducted to obtain an understanding of contemporary dispersal patterns, and hence an understanding of whether hybrid sites have hybrid founders. To avoid confounding effects of history, only relatively newly established sites (determined by a population size of less than eight individuals) were sampled.

Ten variables for every individual in each of 76 populations were measured. These variables were chosen as they differed between parental *Onopordum* species in common glasshouse conditions (unpublished data) and were; plant height, crown width, length of bracts in the outer third series, maximum bract width, floral tube length and corolla length). I also determined the degree that bracts were recurved (0, not recurved; 1, moderately recurved; 2, very recurved), and bract shape (0, expanded base quickly narrowing near base then gradually tapering to tip; 1, tapering evenly from base to tip; and 2, broad base narrowing rapidly only near tip). Finally I calculated a plant shape variable (plant height/crown width), and the degree of branching (number of capitula/plant height).

Distribution of morphological variation

Gower's index (I_G, Gower 1971) was calculated for each continuous character according to the formula:

$$\frac{x_{ij} - x_{j\,\min}}{x_{j\,\max} - x_{j\,\min}}$$

where x represents the value of character j for individual i and min and max represent the minimum and maximum values across all individuals. The mean of I_G for each individual was calculated according to Brochmann I_{hi}(1987), by averaging across all characters. This necessitated the transformation of some values of I_G by 1- I_G so that values for *O. acanthium* were always smaller than *O. illyricum*. This gave a score ranging from 0.0 representing parental *O. acanthium* to 1.0 representing *O. illyricum*. I then characterized each population as the mean hybrid index of each individual within it, and multiplied this value by 100 to get a value between 0 and 100. This gives an index with hybrid populations characterised by intermediate values.

To investigate the distribution of hybrids and parents, these values were categorized into one of five classes (0-20, 21-40, 41-60, 61-80 and 81-100) for plotting on a two-dimensional map of the locations of sites from which samples were collected. Next I examined the way in which the indices changes along major road thoroughfares by plotting the indices obtained from each population against the distance from a starting point for each of arterial roads sampled in this study.

Spatial autocorrelation analysis

Population means and standard errors for each population were calculated for each variable and squared Euclidean distances were calculated between populations for each plant variable. Each variable was treated separately but after standardization between zero and one, several were also combined as; *all characters* (sum of all variables), *combined floral characters* (floral tube length, corolla length, bract length, maximum bract width, degree of recurvature, bract shape). These combinations were produced so that I could obtain multivariate spatial autocorrelation coefficients as this can increase the probability of uncovering spatial structure (Smouse & Peakall, 1999). As all populations occurred near roadsides, and were therefore probably colonized by road-dispersed propagules, spatial distances between populations were calculated by the shortest road distances between sites. Spatial class sizes of 20km were created and spatial autocorrelation was performed using GenAlEx, in which Smouse and Peakall's (1999) spatial autocorrelation for distance matrices can be performed. Significance for these correlograms was calculated by performing 1000 permutations (Peakall & Smouse 1998). While this method was developed specifically for genetic distance matrices, it is a general framework that can accommodate any kind of distance matrix, provided that it is additive and Euclidean. The method is especially useful as it can be used to perform spatial analysis on a matrix that incorporates more than a single locus,

or variable, and hence produces a multivariate spatial autocorrelation which can remove the stochastic noise produced at individual loci/variables.

Sampling of genetic material

Based on the findings of O'Hanlon *et al.* (1999) and *Chapter Six*, I re-sampled seven strategically chosen populations of *Onopordum* for more detailed genetic analysis. These seven populations were chosen to span the range of possible genotypes, and included; one population each of parental species, one that resembled later generation hybrids, and two more each that resembled backcross generations to both parental species (Table 1). Seed was collected from these sites along a 40m transect at intervals of c. 3m. After germination, DNA was extracted from 98 seedlings, representing 14 non-sibs from each site.

Laboratory methods

We used the amplified fragment length polymorphism (AFLP) (Lin & Kuo 1995, Vos *et al.* 1995) assay to generate genetic markers as described in *Chapter Three.* Previously using this method I was able to generate 41 polymorphic markers, 13 of which were diagnostic for *O. acanthium* and 16 for *O. illyricum* (Chapter Five). By having this many segregating markers, it should be possible to distinguish between many different hybridized individuals (Boecklen & Howard, 1997) and allow an accurate estimate of population genetic structure. Primers used for selective amplifications were blue labeled EcoRI-aca primer, 2.25uL of MseI-CTA primer.

Population genetic structure

Scoring of genetic profiles was conducted using an ABI Sequencer and a genetic

distance matrix was constructed as outlined in *Chapter Six.* In order to obtain a graphical representation of the genetic structure of *Onopordum* populations, the pairwise genetic distance matrix was subjected to two dimensional non-metric multidimensional scaling (NMDS, Kruskall 1968) in SYN-TAX (Podani 1995) software package. The subset of the morphological pairwise matrix containing individuals from the same population from which samples were taken for genetic analysis was also subjected to NMDS in an identical fashion. NMDS was performed 30 times from an initial random configuration of individuals in two-dimensional space to minimize the possibility of finding a local rather than global minima. A further iteration was performed using the resemblance matrix from principal coordinate analysis (PCO) as an intital configuration for iteration to test the relative fit of PCO and NMDS to the original distance matrix. Analysis of Molecular Variance (AMOVA) (Excoffier *et al.* 1992, Peakall *et al.* 1995) was also conducted to quantify the degree to which populations were isolated from one another in the GenAlEx (Peakall & Smouse 1998) software package. I also calculated the sum of square deviation values for individual populations, providing an indication of the relative magnitudes of variation within them.

Results

Distribution of morphological variation

Figure 7.1 shows the locations from which plants were sampled as well as the identity of the population in regards to their average hybrid indices. It is clear that the distribution of certain types is not random with similar hybrid classes frequently clustering together. The area west of 148.5°E and north of 35°S is rich in hybrids

intermediate between *O. acanthium* and *O. iilyricum* (black triangles) as well as hybrids resembling *O. illyricum* (diagonal crosses) and relatively poor in plants resembling *O. acanthium*. In contrast, the area south of 36°S contains populations characterized by the presence of plants resembling *O. acanthium* (open squares) or with an affinity toward *O. acanthium* (vertical/horizontal crosses); with a similar pattern for the area north of 35°S and east of 149.2°E. There appears to be two points of transition where areas containing one species rapidly resemble the other. Firstly there appears to be an east-west divide or gradient north of 35°S. On the east plants resemble *O. acanthium* but on the north-west road at c. 148.7°E a mixture of samples is found. West of this point is rich in plants resembling *O. illyricum*. The second boundary (again separating an east-west divide) appears to be found in the diamond shaped area bounded to the west by. 149°E, to the north by 35.3°S and to the south by 35.6°S. While this area is relatively small, there appears a sharp discontinuity between samples resembling *O. acanthium* to the east and *O. illyricum* to the west.

In order to more accurately assess the nature of boundaries between populations I plotted (Fig. 7.2) the actual hybrid index for each population, rather than hybrid classes, against their position along major road routes. Again it is evident that certain areas are rich in populations resembling each other. Often within an area close to 50km transect along a road, the hybrid index of a population does not change markedly. However, it is also evident that sudden changes do occur over very short periods. There appears to be a mixture of gradual and sudden change in hybrid indices along transects (Fig. 7.2).

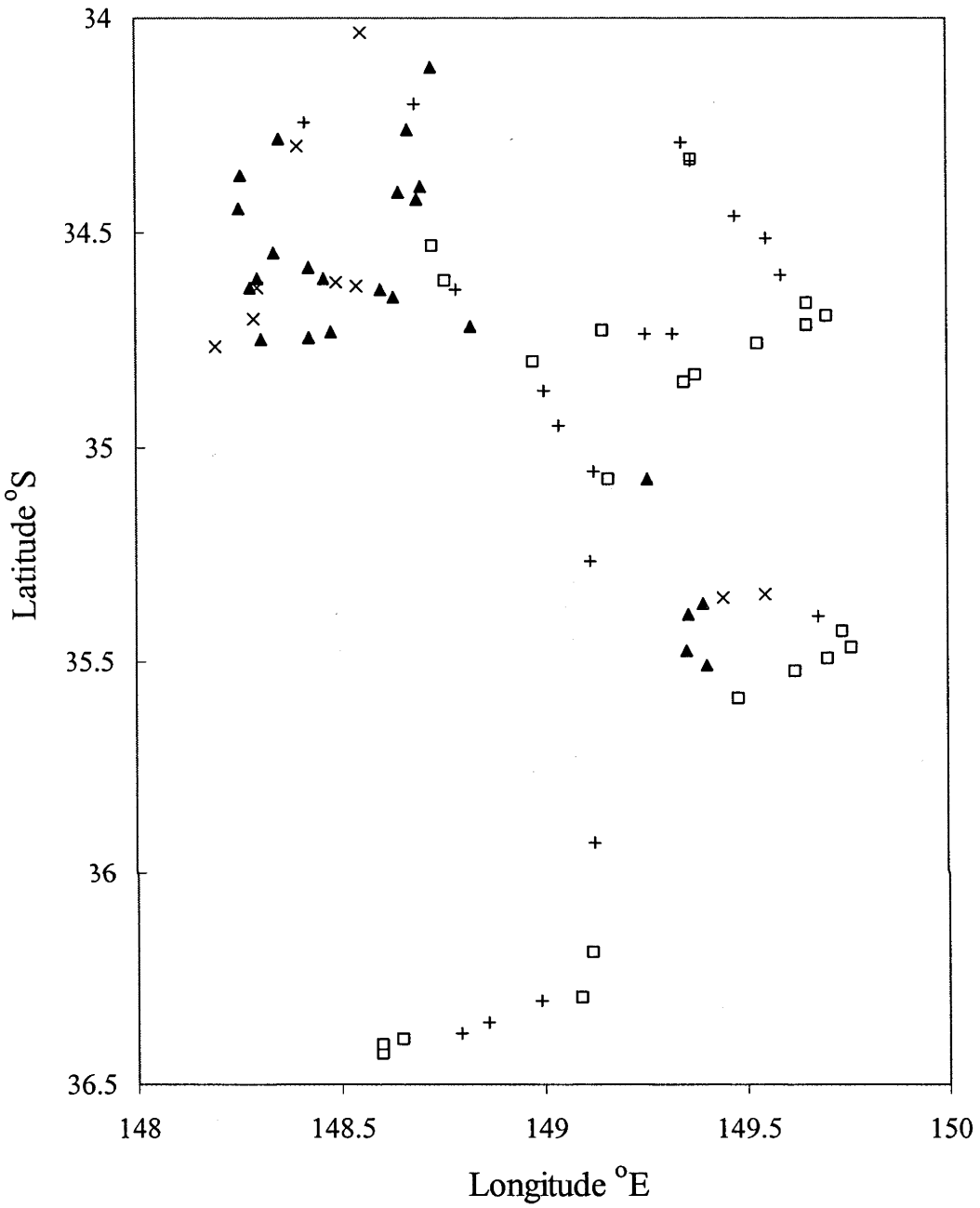


Figure 7.1. Road sites sampled for morphological variation in New South Wales. Samples were collected from newly established sites only, determined by a population size of less than eight plants. Entire populations were sampled. Hybrid indices are classified into 5 classes (open squares - *O. acanthium*; vertical/horizontal crosses - *O. acanthium* affinity; triangles - intermediates; diagonal cross - *O. illyricum* affinity; and closed circle - *O. illyricum*). The only sites containing *O. illyricum* fall outside of the geographic range shown (Maryborough 37°03'S, 143°44'E, Quirindi 31°30'S 150°41'E).

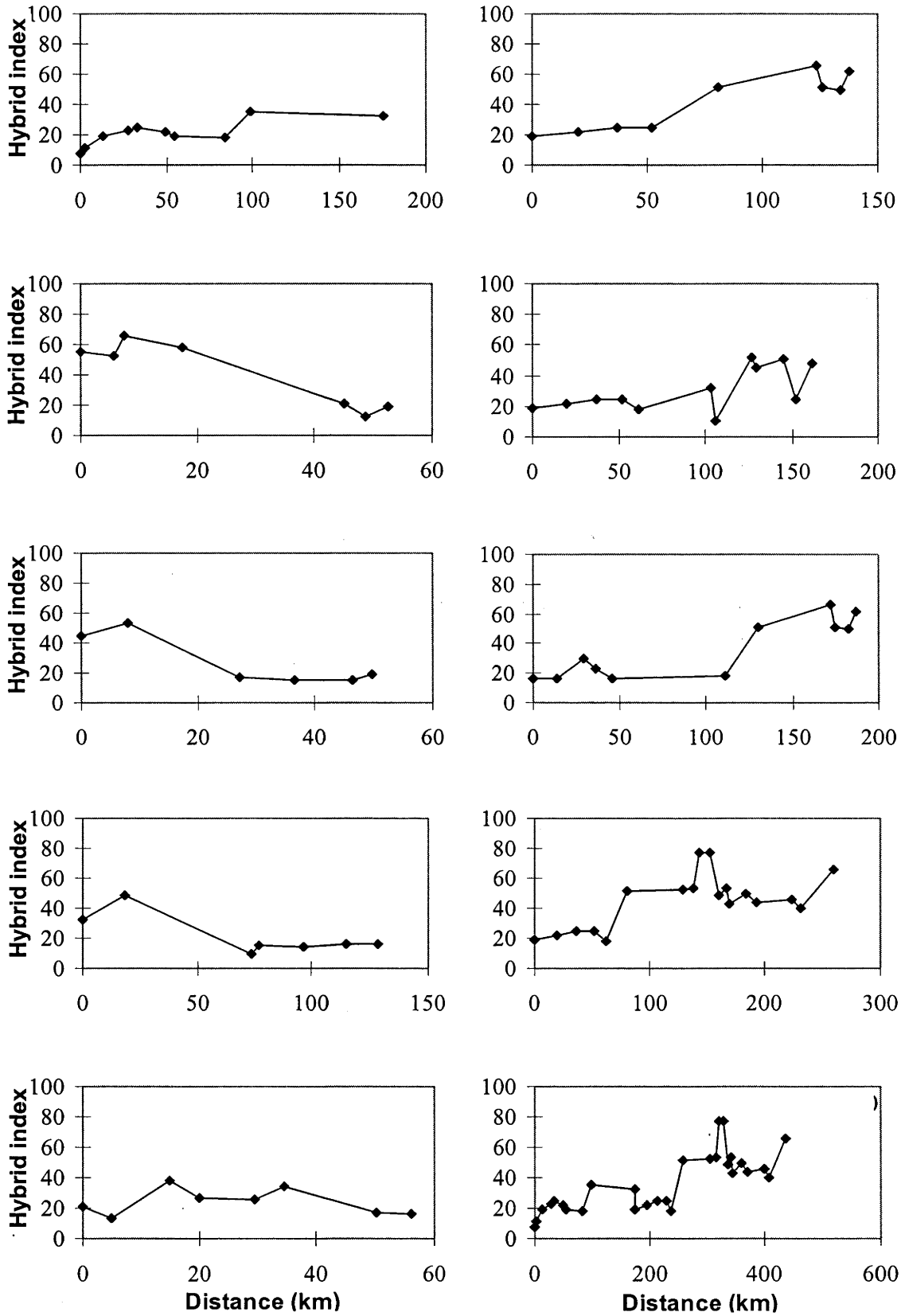


Figure 7.2. Population average hybrid indices along transects in New South Wales.

Spatial autocorrelation analysis

Within the first distance class (20km), I detected very high levels of spatial autocorrelation (Fig. 7.3), especially for the characters relating to floral or capitulum structure. Such autocorrelation gradually declined until 60-80km for most characters, but remained statistically significant to this point (Fig. 7.3). Such a result indicates a high frequency of local gene flow relative to long distance dispersal. As the populations sampled were newly established, it is likely that most of these populations were derived from dispersal events from neighboring sites.

AFLP polymorphism and genetic analysis

Each sample generated an average of 28.4 (standard deviation 4.5) fragments within the scorable size range of 75-450bp. With a single AFLP primer pair combination, a total of 92 fragments were identified, 41 of which were polymorphic. Of these, 29 were diagnostic for either parental species, providing the high resolution sometimes required to distinguish between different classes of hybridized genotypes (Boecklen & Howard 1997, Rieseberg & Linder 1999).

As expected the range of variation observed in this study spanned the full range of diversity between *O. acanthium* and *O. illyricum* (Fig. 7.4). NMDS of the genetic distance matrix produced a plot with a stress value of 13.0% deviation between the final NMDS result, and the original distance matrix (Fig. 7.4). Such a value compared favorably with that produced by PCO, which showed a poorer fit to the data (Stress = 19.3%). A similar pattern of stress was obtained from the morphological data (NMDS Stress 2.7%, PCO Stress = 10.5%, Fig. 7.4B). In both morphological and genetic analyses, populations of the parents, *O. acanthium* and *O. illyricum* were characterized by low levels of genetic diversity, as evident in the limited spread of genotypes in these populations (Fig. 7.4). In contrast, populations derived through

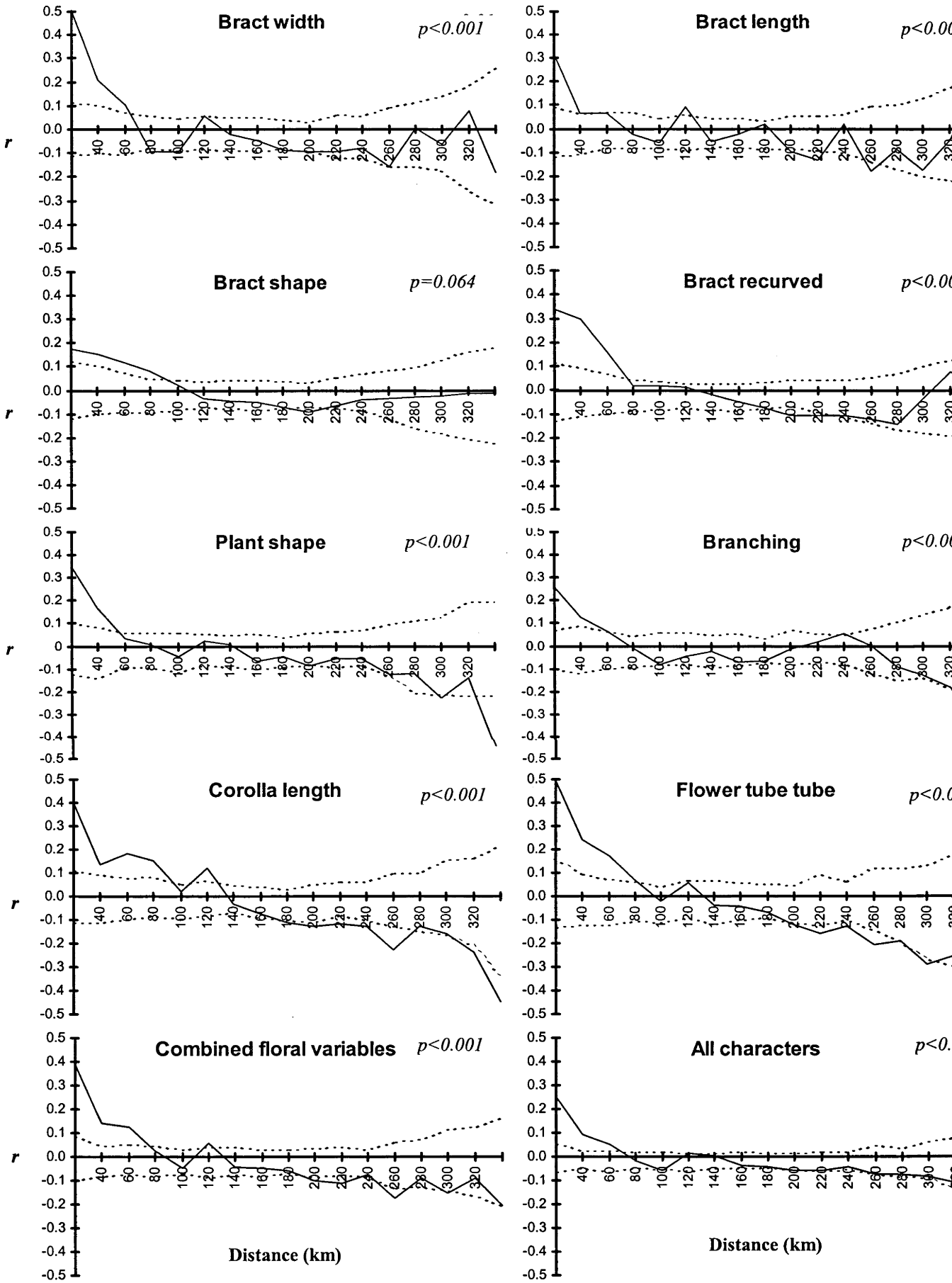


Figure 7.3. Correlograms (solid lines) for each morphological variable analyzed separately, and the corresponding combined multivariate correlograms, with 95% null hypothesis confidence regions indicated by dotted lines; the statistical significance of total correlograms are shown in the top right corner.

hybridization appeared to contain greater genetic diversity, although no single hybrid population spanned the full range of diversity between *O. acanthium* and *O. illyricum* (Table 7.1, Fig. 7.4). For genetic and morphological data, the infestation at Bungendore (black diamonds, Fig. 7.4) appeared to contain the greatest spread of genotypes ranging from intermediacy to types similar to *O. illyricum*, but the population at Braidwood (black triangles, Fig. 7.4) clustered tightly near *O. acanthium*. All populations previously defined as being of hybrid origin by two samples contained no genotypes that were characteristic of either one of the parental species.

Both of these analyses gave similar patterns of population structure, though in the morphological analysis (Fig. 7.4B), few individuals appeared to fall near the mid-point between parents, perhaps due to unpredictable patterns of inheritance (Rieseberg & Ellstrand, 1993).

Analysis of Molecular Variance (AMOVA) confirmed the patterns revealed by NMDS. 70% of the genetic variation was partitioned between populations (Table 7.1). Pairwise ϕ_{st} values show that the all the populations were significantly differentiated from each other. Hybrid populations contained much higher levels of diversity than either of the populations representing parental species (Table 7.1). The magnitude of variation present in populations differed between hybrid populations. Braidwood was the site containing the lowest level of variation while three times this level was observed in Bungendore.

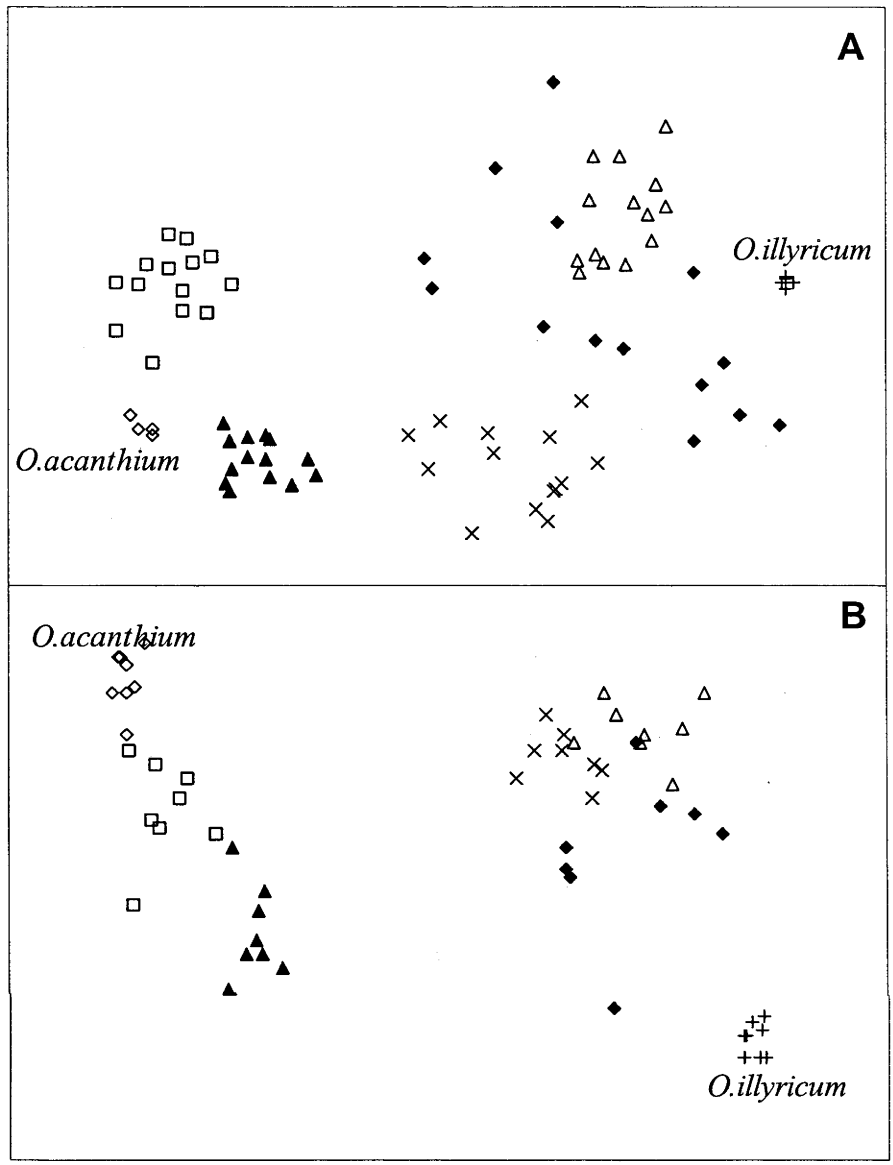


Figure 7.4. Non-metric multidimensional scaling plots of AFLP variation (A) and morphological variation (B). Sites are represented by (Orange - open diamonds; Pialligo - open squares; Braidwood - black triangles; Galong - open triangles; Captains Flat - diagonal crosses; Bungendore - black diamonds; Maryborough - plus symbols).

Table 7.1 Analysis of Molecular Variance of AFLP data for seven populations of *Onopordum* spp. with f_{st} as an analogue to F_{st}.

	d.f.	S.S.	M.S.	Variance %	ϕ_{st}	*p*
Source						
Among Pops.	6	520.551	86.759	70		
Within Pops.	91	232.037	2.549	30	0.703	<0.001
Populations				latitude °S	longitude °E	
Orange	13	1.714		33°17'	149°06'	
Pialligo	13	43.571		35°17'	149°13'	
Braidwood	13	26.929		35°27'	149°48'	
Galong	13	38.714		34°36'	148°34'	
Captains Flat	13	42.286		35°35'	149°27'	
Bungendore	13	77.929		35°15'	149°27'	
Maryborough	13	0.894		37°03'	143°44'	

Discussion

Spatial analysis of morphological variation

As a colonizing weed with limited wind dispersal (O'Hanlon *et al.* 1999; *Chapter Six*), *Onopordum* infestations are patchily distributed. Even in densely infested areas, *Onopordum* generally exists as patches, probably brought about through very local dispersal. This pattern also appears to be repeated at a higher geographic scale. For broader structure in the distribution of variants to develop in this group (Figs 7.1 & 7.2), either (i) dispersal over a large scale has been non-random or (ii) environmental variation has selected for such a structure. Elsewhere (*Chapter Eight*) I examine the influence of bioclimatic selection, but here I limit my discussion to patterns of dispersal and founder events.

Analysis of the distribution of population hybrid indices (Figs 7.1 & 7.2) revealed that large areas contained populations composed of plants of similar identity within the spectrum separating *O. acanthium* and *O. illyricum*. Spatial autocorrelation analysis revealed extreme spatial autocorrelation within distances of 20km (Fig. 7.3) indicating that local dispersal was important in the distribution of morphological variation. In addition, for most characters significant autocorrelation was observed between populations up to 40 and 60km apart. Beyond this distance, there was no significant overall relationship between the identity of individuals composing populations. However, several long transects taken individually through this range indicated that the identity of plants did not change substantially up to a distance of 100km in some cases. From these analyses, I conclude that local dispersal is an important factor influencing the contemporary structure of *Onopordum* populations, producing a

patchy distribution with sites being established by migration from other nearby sites. Since dispersal occured mostly at a local scale, it is reasonable to infer that many hybrid sites were established through secondary dispersal from a population containing hybrid plants rather than primary contact between parental species. A similar situation was shown for hybridizing *Carduus* thistles (Warwick *et al.*, 1989).

However, such patchiness is occasionally interrupted by populations of very different identity, probably as a consequence of long distance dispersal events. Figure 7.2 shows several examples of population identities changing dramatically within a short distance (<5km). Unlike many recorded examples of hybridizing taxa, at least one of the *Onopordum* species studied exhibit a mixed mating system (Michalakis *et al.* 1993) and capitula from *O. acanthium*, *O. illyricum* and hybrids that were forced to self also set viable seed (pers. obs.). It is therefore possible that populations may be founded by single migration events of few, or even single propagules. The opportunity for a long distance dispersal into a location that generally is characteristic of a different variety of *Onopordum* may therefore be high. As *Onopordum* can set a vast number of seed and has such a large seed bank (Pettit *et al.* 1996), once the identity of a population becomes established over the first few generations, it is unlikely that subsequent and infrequent migration would change its identity without strong selection. Consequently, neighboring populations of different genetic makeup could persist with minimal effect on each other and a region may contain many similar populations derived from local dispersal, but also several different sites derived from outside of the area. Through a combination of long and short distance dispersal, many divergent populations consisting entirely of hybridized genotypes could have become established. Such an argument is also supported by more detailed within population genetic analyses.

Limited variation in populations

Given that morphological data supports the idea that hybrid sites are most often established from neighboring hybrid sites, one may expect that hybrid sites contain relatively narrow subsets of the complex's overall variation. Both the AMOVA and NMDS showed that some hybrid populations of *Onopordum* contained very low levels of genetic diversity (Table 7.1, Fig. 7.2). However, as expected the hybrid populations studied here were more variable than those containing parental species (Table 7.1, Fig. 7.2). Such patterns are consistent with effective isolation of hybrid populations and little between-site dispersal, though not all hybrid populations were equally diverse. This is consistent with initial hybridization events that produced the various progenitors of the populations sampled here. Populations such as those at Bungendore (black triangles, Fig. 7.2 and Table 7.1) may be more diverse as a consequence of either the merging of spatially adjoining but genetically differentiated populations, or by several different genotypes founding the infestation. Furthermore, hybrid populations were each differentiated from one another (Fig. 7.2), inconsistent with expectations of *in-situ* hybridization or colonization from the same original hybrid population.

Selection is another factor that may have acted over time to produce hybrid populations with limited genetic variation. Groves *et al.* (1990) in studying allozyme variation at different life-history stages have shown that in *Onopordum*, selection operates strongly in the establishment and recruitment of an *Onopordum* cohort. In that particular population studied (Galong; open triangles Fig. 7.2: shown here to be a hybrid lineage), heterozygosity increased within the cohort over time. Such a finding

may be taken as evidence of later generation heterosis in *Onopordum* hybrids. However, given that all the hybrid populations studied here had diverged from one another, it seems likely that other processes such as historical ones have contributed more to the local contemporary patterns.

Consequences for hybrid speciation

Much attention has focused on mechanisms for hybrid speciation in sympatry, but the example of invasive *Onopordum* thistles illustrates that speciation following the isolation of hybrids is possible during the colonization of new localities, and therefore of potential interest for understanding speciation. Despite the difficulties that hybrids may encounter in surviving the 'challenges of sympatry' with parents (Rieseberg and Carney 1998), the establishment of allopatric populations eliminates such challenges, and may have assisted in the establishment of isolated hybrid *Onopordum* lineages. Whether hybrids have a fitness advantage in the Australian context and whether a degree of reproductive isolation contributes to the maintenance of hybrid distinctiveness is unclear. Such questions are important, and may be essential in determining whether hybrids can survive future contact with parents in the long term. More broadly, while examples of isolation of hybrids from parents are few, the semi-natural setting of the colonization by weedy *Onopordum* of Australia may be analogous to the hybrid invasion of new localities after changes in sea level or glacial retreat. While such events are sporadic, they may represent occasions during which allopatric hybrid speciation is an important process, especially for organisms capable of long distance dispersal and which are self compatible.

Chapter Eight *Identification of a narrow hybrid zone between* Onopordum *species indicates ecologically-dependent speciation*

The fate of hybrids in different ecological contexts can provide important information on processes related to speciation or coalescence between lineages. Elsewhere (*Chapter Nine)* hybrids between Scotch Thistle (*O. acanthium*) and the Illyrian thistle (*O. illyricum*) are shown to be intrinsically no less fit than their parents in glasshouse conditions. Furthermore, hybrids have become widespread in Australia (part of their alien range). However, there is evidence to suggest that hybrid zones are more transitory and less extensive in their native European range. Here I (i) identify a hybrid zone in Europe, (ii) show that it is narrow, occurring at an intermediate elevational range (iii) contrast the pattern with that observed in Australia, (iv) argue that selection against hybrids is the critical force maintaining species boundaries in Europe and (v) suggest that the abundance of hybrids in Australia represents a relaxation of such destabilizing selection, with a concomitant breakdown in species boundaries.

There are no uniform evolutionary consequences of natural hybridization (Arnold & Hodges 1995). However, the maintenance of hybrid zones represents a special scenario whereby species maintain their separate identities in the face of ongoing inter-specific gene flow (sympatric speciation). Various models have been developed in an attempt to explain hybrid zone maintenance (see review by Rieseberg 1999 and references within). Despite early work showing the importance of habitat variation to hybrid fitness, these models have historically proposed that the fitness of hybrids was genetically determined and invariant with habitat (e.g. Barton & Hewitt 1985). Such a model predicts that selection against hybrids and recurrent hybrid formation produce a stable hybrid zone. More recently the influence of environmental variation in determining the fate of a hybrid zone has again been demonstrated (Emms & Arnold 1997). Models incorporating such evidence usually predict that hybrids are selected

against in parental habitats but parental species are selected against in a 'hybrid habitat'. However, that the fitness of hybrids is variable across environments does not also preclude the possibility that hybrids also have some element of genetically determined hybrid inferiority (Bert & Arnold 1995) and in nature, the maintenance of hybrid zones may be extremely complex.

An understanding of the ecological mechanisms underlying the dynamics of hybrid zones may be best achieved when hybrid zones can be manipulated, or using a comparative approach where hybrid zones have different fates in different ecological conditions. Such studies are difficult to perform and attempts to do so have usually taken the form of comparisons of hybrid fitness in controlled and natural conditions (e.g. Hatfield & Schluter 1999), or in studying the relative fitness of hybrids and parental species in their repective habitats (e.g. Emms & Arnold 1997, Wang *et al.* 1997, Burke *et al.* 1998, Maccallum *et al.* 1998). An alternate approach may be in the examination of the different fates of hybrid zones when species that hybridize in their native range expand their ranges by invading ecologically novel localities, or when the environment is altered (see examples in Anderson 1949, Rieseberg 1999). Such analyses may reveal which habitat parameters are important in the maintenance of hybrids zones (and therefore speciation) or in the homogenization of lineages.

Onopordum is a Eurasian thistle genus, species of which have been introduced into several continents including Australia. Interestingly, hybrid zones in the different continents appear to have different evolutionary fates. In Australia the presence of hybrids between *O. acanthium* and *O. illyricum* have been confirmed by genetic analysis (O'Hanlon *et al.* 1999). Subsequent more detailed genetic analysis has

revealed that most infestations are composed entirely of hybridized genotypes (*Chapter Seven*), probably as a consequence of the combination of colonization from neighboring sites, and also through selection favoring hybrids in certain bioclimatic conditions. In addition, in *Chapter Nine* I show that in glasshouse conditions, hybrids are not intrinsically unfit relative to their parents. However, in Europe the ranges of *Onopordum* species are often allopatric but come into contact with other *Onopordum* species in relatively low densities in a few locations (Franco 1976, Briese 1990). While hybrids appear to be frequently formed wherever parental species come into contact, anecdotal evidence suggests that they never become abundant at such sites, leading to the formation of a narrow hybrid zone. Furthermore there are no records of hybrids existing in the absence of parents, and hybrids may not introgress into parental populations. These factors suggest that selection operates against hybrids in their native range, and that they exist only through recurrent hybridization between parents. However, further work documenting the existence of such a narrow hybrid zones, and the quantification of associated environmental variation in the native range is required before the mechanisms of hybrid zone maintenance can be determined.

In this chapter, my specific objectives are to (i) describe the morphological variation in a European putative hybrid zone, (ii) determine the width of any confirmed hybrid zone, and (iii) relate its location with ecological conditions or elevation. Finally, these analyses provide evidence that selection against hybrids is the critical force maintaining species boundaries in Europe and that the abundance of hybrids in Australia represents a relaxation of such reinforcing selection, with a concomitant breakdown in species boundaries.

Materials and Methods

To examine gene flow between populations, I surveyed morphological variation in all populations of *Onopordum* neighboring Montpellier, France. Plants from 14 populations were harvested. These populations were putatively classified into those comprising pure *O.acanthium*, *O. illyricum* or containing at least some intermediate individuals (Table 8.1). These populations could be placed along a 70km transect running from the coast inland, so that sites were ordered from low (277m) to high (824m) elevation (Table 8.1). Such an elevantional change corrresponds with a bioclimatic gradient between the Mediterannean coastal climate and a cooler temperate zone at higher elevations. The distances between sites was measured along the transect for the graphical purposes of examining the distribution of morphological variation in relation to distance and elevation (Fig. 8.2).

Table 8.1 Details of the locations, elevations, putative identifications based on morphological inspection and the number of samples analyzed for populations of *Onopordum* near Montpellier (Herault), France.

Location	Longitude(E)	Latitiude(N)	Putative identity	Elev.(m)	*n*
Cazavielle	3°47.45'	43°16.38'	*O. illyricum*	277	21
Viols-en-Laval	3°43.86'	43°14.78'	*O. illyricum*	255	16
les Matelettes	3°40.94'	43°13.24'	*O. illyricum*	286	12
Ferussac 1	3°30.02'	43°17.58'	*O. illyricum*	618	10
Ferussac 2	3°29.98'	43°17.56'	*O. illyricum*	624	15
Ferussac 3	3°29.93'	43°17.51'	*O. illyricum*	626	11
Mas-de-Jourdes	3°28.64'	43°19.81'	*O. illyricum*	579	8
la Trivalle 1	3°29.08'	43°16.22'	*O. illyricum*	660	11
la Trivalle 2	3°29.06'	43°16.22'	*O. illyricum X acanthium*	670	27
la Trivalle 3	3°29.05'	43°16.22'	*O. illyricum X acanthium*	673	16
le Caylar	3°19.44'	43°21.83'	*O. acanthium*	808	13
la Salvetat	3°15.87'	43°24.86'	*O. acanthium*	768	11
la Blaquererie	3°13.51'	43°27.06'	*O. acanthium*	754	11
la Cavalerie	3°09.69'	43°30.84'	*O. acanthium*	824	10

Six floral traits were measured from all plants within each population (Table 8.2). This was achieved by digitally scanning prepared samples and importing images into the software package Sigma Scan (SPSS. Inc) for measurement. Characters measured were bract length, bract width at the widest point, bract shape (which was calculated by the formula 4.Π.Area/Perimeter giving a value ranging between 0 for a circle and 1 for a straight line), corolla length and floral tube length. For each character, an analysis of variance was performed after testing for normality and the equality of variances. Where these tests failed, a Kruskal-Wallis analysis of variance on ranks was performed (Fig. 8.2). All statistical analyses were performed using the software package Sigma Stat (SPSS. Inc).

Kramer values for each variable were calculated using the PATN (Belbin 1983) software package. These values are the proportion of variation between diagnostic groups relative to the total variation, giving a score between 0 and 1. A score of 1 is achieved when all the variation for that character is partitioned among the groups that one is trying to differentiate while 0 indicates that the character gives no diagnostic power, with all the variation being partitioned within groups. Values of greater value than 0.7 generally indicate that a character is diagnostic, and so all characters generating values greater than this were included in additional analyses (Belbin 1983). In order to calculate these scores, the three most extreme populations at the ends of the transect were chosen as the groups to differentiate as these were most likely to represent parental *O. acanthium* and *O. illyricum*. Characters determined to be useful under those criteria were then used for ordination by principal components with correlation (Fig. 8.1), with inter-individual distances calculated by the Gower Metric (Belbin 1983).

Finally, hybrid indices were calculated for all individuals using Gower's index (I_G, Gower 1971) for each continuous character as of *Chapter Seven*. This gave a score ranging from 0.0 representing parental *O. acanthium* to 1.0 representing *O. illyricum*. Population averages were calculated and plotted against the distance along the transect and against elevation. Population averages and standard errors for individual characters were also computed and plotted similarly. This enabled a visualization of the shape of the hybrid zone.

Results

A range of Kramer values was produced for the characters analyzed in this study. A value of greater than 0.7 usually indicates that the character can distinguish between different groups (Belbin, 1983). Here, all the characters analyzed except corolla length were found to be diagnostic (Table 8.2) on this basis. Therefore corolla length was excluded from subsequent analyses. Bract width and bract shape in particular were notably diagnostic and had Kramer values of greater than 0.9 (Table 8.2). Therefore, for this study, five characters that were each diagnostic for the different species were identified.

Table 8.2 Kramer values for each of the characters analyzed.

Variable	Kramer value
Bract length	0.713108
Bract width	0.943300
Bract shape	0.904095
Corolla length	0.408661
Floral tube length	0.725967

Principal component analysis revealed that species were discrete, separated by an axis that accounted for 81.12% of the total variation (Fig. 8.1). In addition to parental species, there was also a high frequency of intermediate samples identified on the scatterplot. Individuals from populations containing hybrid individuals are represented by diagonal crosses (Fig. 8.1). These intermediates were sampled only from populations described as containing putative hybrids based upon field inspection (Table 8.1). In addition to hybrid individuals, these populations also appeared to

contain individuals corresponding to both parental species (Fig. 8.1). Figure 8.1 reveals not only a large spread of variation among species, but also within them, especially along the vertical axis. However, this axis accounts for only 8.74% of the total variation and therefore as the axes are not to scale, the degree of variation in this direction is exaggerated. All other dimensions accounted for less than 5% of the variation and are therefore not shown.

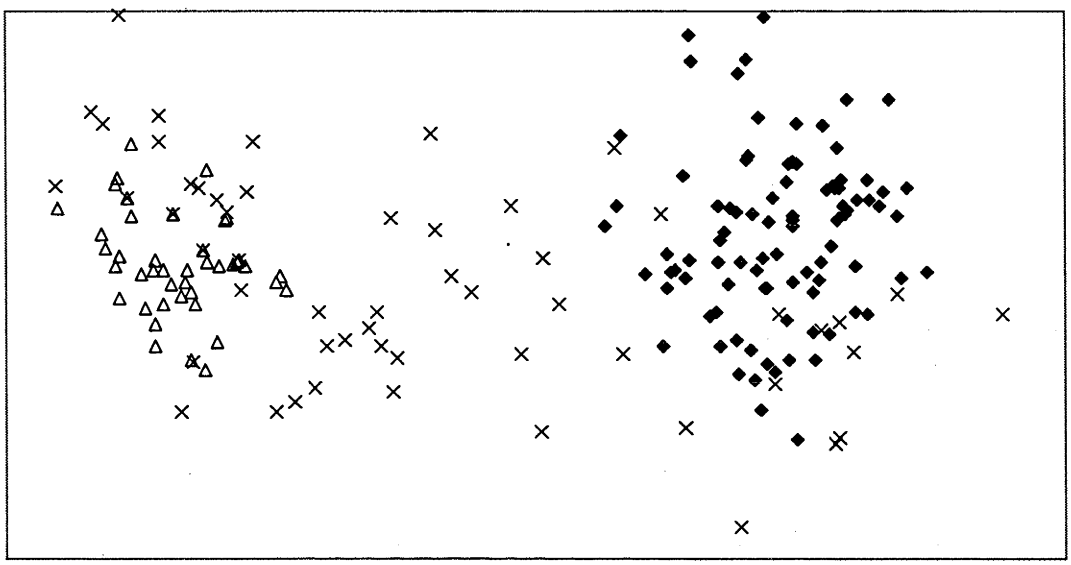


Figure 8.1 Principal coordinate analysis of all individuals grouped into three classes correposnding to the putative population identities in table 8.1 (Open triangle - *O. acanthium*, diagonal crosses – populations containing intermediates, closed diamonds - *O. illyricum*). The horizontal axis for 81.12% of the total variation and the vertical for 8.74% (Total 89.96%).

The elevations of the populations sampled ranged between 220m and 824m. For each character studied, plants at higher elevations had lower values than those occuring on the coastal plane (Fig. 8.2). Analyses of variance showed that populations differed significantly from each other for each variable studied (Fig. 8.2). There was a correspondence between the putative identity of populations and the degree of differentiation between them. Populations identified as those comprising either of the two parental sites produced the most extreme values for each character, while sites containing putative hybrids had an average value that was intermediate. An analysis of the distribution of variation across populations revealed a sharp discontinuity between values that corresponded with the hybrid sites, and also corresponded to an intermediate elevation (Figure 8.2). As a summary, the average hybrid index was relatively uniform across putative parental populations, and was intermediate at

putative hybrid sites. Such a sharp discontinuity in morphological characters indicates a restriction of gene flow between populations, even over relatively short distances.

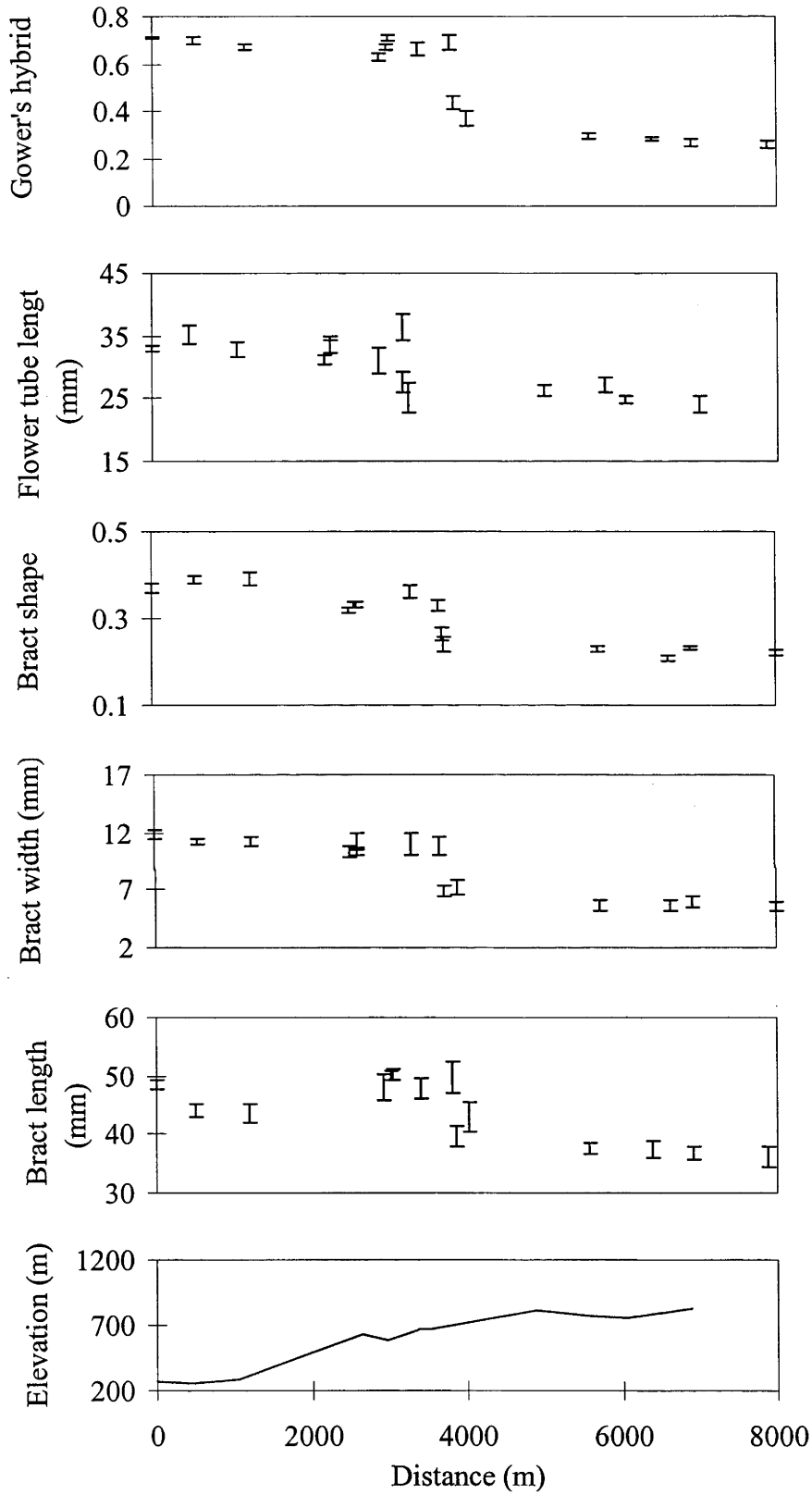


Figure 8.2 Population averages and standard errors calculated for each variable at each site along a transect separating high and low altitude *Onopordum*. The elevation for each site is also shown for comparison with morphological values. High and low altitude populations were distinct for each variable, and sites at intermediate elevtaion were morphologically intermediate.

Discussion

Morphological analyses have revealed that many individuals that were sampled from the putative hybrid populations were of hybrid origin. However, intermediacy alone does not guarantee a hybrid origin (Rieseberg 1997) as other factors such as plesiomorphic intermediacy, convergence, developmental differences, or non-additive inheritance can also give rise to apparent hybridization. However, (i) hybrids between these species have previously been described in Europe, (ii) several intermediate characters were identified, reducing the chance of convergence, (iii) the intermediates correspond in morphology to the hybrids confirmed by genetic analysis to exist in Australia, and (iv) parents co-occurred with hybrids, thus reducing the chance of phenotypic plasticity and developmental factors producing intermediacy. In addition, in *Chapter Six*, a strong correlation between morphological variation and genetic variation was identified. Finally as putative hybrids were identified only in populations also containing parents, it is unlikely that these samples represented a third intermediate species. Therefore the most probable explanation is that individuals were descended from hybridization events between *O. acanthium* and *O. illyricum*.

While hybrid zones between these species have been identified in Australia (O'Hanlon *et al.* 1999), Spain (Gonzalez Sierra *et al.* 1992), Greece (A. Sheppard pers. comm.) and Turkey (P. Hein pers. comm), they have not previously been identified in Southern France, perhaps due to their narrow and possibly transitory nature. In contrast to Australian populations of hybrids, those identified here also contained individuals corresponding to parental species (Fig. 8.1). In addition, the hybrid sites have at least one population comprising 'pure' parental species in very

close proximity (*O. illyricum* at Ferussac and Mas-de-Jourdes) with *O. acanthium* occurring further away (except for in the hybrid site, Fig. 8.2). The sites comprising hybrids and parental species could potentially represent a point of overlap between the species' ranges; at the southern end for *O. acanthium* and the northern end for *O. illyricum* (though in isolated and disturbed sites on the coastal plain *O. acanthium* can occasionally be found, often along with weedy *O. tauricum*; pers. obs.). That the species' ranges appeared to overlap and that hybrids occured only at this point demonstrates that the hybrid zone is particularly narrow, considering that high values of spatial autocorrelation at 60km were observed in Australia (*Chapter Seven*). As the hybrid zone is so narrow and gene flow across the hybrid zone could easily be accomplished, it is apparent that selection has operated to maintain species boundaries. Furthermore, that the hybrid zone appears to correspond with a unique elevation is suggestive that habitat plays an important part in the maintenance of this hybrid zone, and indicates that habitat-dependent selection was probably important in their differentiation and ultimate speciation. Furthermore, the contrasting fates of hybridization between continents, and the relatively high intrinsic fitness of hybrid genotypes (*Chapter Nine*) provides strong support for this hypothesis. While speciation appears to have been ecologically driven in Europe, the change in habitat accompanying the invasion of Australia appears to have resulted in coalescence between these lineages.

Such ecologically driven speciation may have important implications for the evolution of *Onopordum*. As the mechanism for determining divergence or coalescence in *Onopordum* appears to have a large ecological component, the evolution of *Onopordum* species may be a very plastic one, tracking environmental fluctuations. In

particular, episodes of glaciation may have seen *Onopordum* species oscillate between distinct and homogenized lineages as suggested by a phylogeographic analysis of European *Onopordum* species (*Chapter Five*). Such oscillations and the lack of reproductive barriers between species may have enabled *Onopordum* to maintain the phenotypic diversity required to inhabit a broad range of conditions. Such a strategy may have assisted during periods of range expansion, where adaptive traits could be sorted according to habitat variation. The new ecological conditions experienced after colonizing Australia and the mixture of genes from different lineages may have allowed for the subsequent sorting of genes, and selection for hybrids in this unique habitat.

Chapter Nine *Fitness consequences of hybridization in Onopordum and fruit mass variation in hybridizing species*

Several key components of fitness were investigated to explore the fitness consequences of hybridization between *Onopordum acanthium* and *O. illyricum*. These components were (i) achene mass, (ii) achene number, (iii) proportion of unfilled pericarps, (iv) time to germination (v) proportion of germinating seed, (vi) seedling growth rate and (vii) inter-specific competition during early vegetative growth. Large differences were observed between these species for each of these components in uniform conditions. Generally, hybrid derivatives were intermediate between parental species, though, with a more variable value. Specifically for the components (i, ii, iv, vi and vii) hybrids encompassed the entire range of fitness responses, with both the minimum and the maximum response. In line with work in other species, later generation hybrids studied here were not uniformly fit, with certain individuals showing high fitness responses in certain traits. While it was not possible to determine the overall fitness for individuals, for individual responses it was shown that hybrids were not intrinsically unfit, and that given appropriate ecological conditions, certain individuals may have higher fitness than parental genotypes. Hybridization between parental species with different achene masses resulted in a hybrid lineage with a new, intermediate and highly variable achene mass. Partitioning achene mass variation revealed that maternal genotype was critical in determining achene mass within species. The generation of novel achene masses through hybridization has not previously been documented and may be important for species invading ecologically novel environments.

The creative role of hybridization in the evolution of plant species depends on the capacity of hybrids to reproduce beyond several generations. Several factors contribute to this capacity. Firstly, there are pre-zygotic factors, such as spatial constraints, phenology, pollen tube growth and inter-species pollen competition (Arnold & Hodges 1995, Emms *et al.* 1996, Campbell *et al.* 1998, Hauser Shaw *et al.* 1998, Rieseberg & Carney1998), and secondly, post-zygotic mechanisms, such as differential embryo abortion, germination, and plant growth are also critical to hybrid establishment and persistence (Arnold & Hodges 1995, Emms *et al.*1998, Cummings *et al.* 1999, Wang *et al.* 1999). Apart from the intrinsic genetic (endogenous) fitness

of hybrids, many of the pre and post-zygotic factors are dependent on exogenous environmental variation, and therefore habitat variation constitutes a third major factor (Emms & Arnold 1997, MacCallum *et al.* 1998, Wilhelm & Hilbish 1998, Hatfield & Schluter 1999).

The ecological consequences of hybridization in plants have been explicitly studied in hybrids for nearly a century (see reviews by Anderson 1949, Rieseberg & Carney 1998). The accumulated data have revealed no uniform fitness consequence of hybridization between plant species. Firstly, in terms of endogenous fitness, consequences ranging from genetic incompatibility and incomplete zygote formation or sterility to heterosis have been observed (Arnold & Hodges 1995, Rieseberg 1997). Indeed over several generations, the intrinsic fitness of hybrids and their derivatives may change (Rieseberg *et al.* 1996, Rieseberg 1996). Mechanisms proposed to explain hybrid vigor are dominance, overdominance, and epistasis (Hedgecock *et al.* 1995, Edmands 1999, Fritz *et al.* 1999) allowing beneficial genes from different parents to be combined and potentially interact in hybrids. When heterosis occurs, it is often followed by a reduction of fitness in later generations, through the segregation of co-adapted complexes (Edmands 1999, Fritz *et al.* 1999). The reverse has also been observed with F_1 offspring displaying mild sterility, but a return to full fitness after several generations and backcrossing (Rieseberg *et al.* 1996, Arnold *et al.* 1999). Where hybrids are intrinsically fit enough to persist, exogenous factors have also been observed to influence overall fitness. The relative fitness for many traits relating to reproductive and vegetative growth have been shown to be under environmental control, whether biotic, such as increased achene herbivory in hybrid sunflowers (Cummings *et al.* 1999), inter-species pollen competition in sunflowers (Rieseberg *et*

al. 1998) and plant community structure in Louisiana irises (Emms & Arnold 1997) or abiotic, such as light availability, soil moisture and soil chemistry in Louisiana irises (Cruzan & Arnold 1993, Emms & Arnold 1997) and chemical uptake in big sagebrush (Wang *et al.* 1999).

In this chapter I explore a range of fitness components at different life-history stages in hybrid *Onopordum*, focusing on the intrinsic fitness or 'unfitness' of hybrids without reference to environmental variation. It has previously been shown that *O. acanthium* (Scotch or Cotton thistle) and *O. illyricum* (Illyrian thistle) hybridize in both their native European range (Gonzalez-Sierra 1992; see also *Chapter Five*), and in their alien range (O'Hanlon *et al.* 1999; see also *Chapters Six* and *Seven*) where they are weeds. While hybrids occur frequently in Europe, they appear to represent a transitory group, and never exist in the absence of parental species (Gonzalez-Sierra 1992). In contrast, in Australia, hybrids represent the most abundant form in the complex with parental species being less common than hybrids (O'Hanlon *et al.* 1999; *Chapter Seven*). In Europe strong climatic preferences appear to separate the species, *O. acanthium* being present mostly in more highly elevated and temperate conditions than *O. illyricum,* which occurs on the coastal Mediterranean fringe (Franco 1976; *Chapter Eight*). Such a distinct boundary is not apparent in Australia, and therefore suggests an important role for environmental variation in the outcome of hybridization. However, for exogenous factors to be important, hybrids must be intrinsically fit enough to reproduce over several generations. As the relative fitness of hybrids is unknown, an understanding of the relative fitness of hybrids at various life history stages will allow a greater appreciation of the consequences of hybridization in different environments.

Of the various fitness components studied here, achene mass variation deserves special attention. Life-history theory predicts that organisms should produce offspring of a single optimal size (Smith & Fretwell 1974, Haig & Westoby 1988). In plants, seed size and to a lesser degree fruit size is considered one of the least plastic characters (Harper *et al.* 1970, Fenner 1985), because the optimal achene mass (seed and fruit in *Onopordum*) will maximize the return per unit investment, resulting in a trade-off between achene size and achene number (achenes are single seeded fruit and hence the same return on unit investment applies). Nevertheless, in recent years offspring size variation has been reported in many species, conflicting with life history theory (Thompson 1984, Vaughton & Ramsey 1997) and it is now well established that many different types of plants exhibit intra-specific seed mass variation. Several explanations for this variation exist, ranging from developmental, resource, and/or spatial constraints (Wolfe 1995, Vaughton & Ramsey 1997) to various genetically determined mechanisms selected for through different dispersal mechanisms, or germination times (Venable & Brown 1988, Geritz 1995), by interactions between plant size and reproductive behaviour leading to assortative mating (Venable 1992, Sakai & Sakai 1995), interactions with herbivores (Moegenburg 1996), through preferential allocation of parental resources to high-quality genotypes (Temme 1986, Waser 1993) or hormonal control of the parent by the developing zygote (Lipow & Wyatt, 1999).

While offspring mass variation has been the focus for study within species, the consequences of hybridization between plants with different offspring size remains unknown. Potentially, hybrids may demonstrate greater offspring mass variation than parental species, providing scope in segregating hybrids to investigate the genetic

basis for offspring mass variation. Whether the variation is partitioned equally among highly variable plants, or whether offspring mass is constrained within individual plants and whether hybrid offspring have a unique mass will provide key insights into the mechanisms of offspring mass variation. Indeed a demonstration of a stable shift in offspring mass accompanying hybridization would provide considerable support to the adaptive significance of offspring mass variation. As a first step in understanding how offspring mass variation is influenced by hybridization, I examined the natural achene mass variation in the hybrid complex of self-compatible biennial thistles in the genus *Onopordum*.

Materials and Methods

Study Species

Onopordum is a Eurasian thistle genus, species of which have been introduced into several continents including Australia where they have become major pasture weeds. In Australia the presence of hybrids between *O. acanthium* and *O. illyricum* have been confirmed (O'Hanlon *et al.* 1999, *Chapters Six* and *Seven*). They are monocarpic and have life cycles ranging from 1 to 7 years (Pettit *et al.* 1996). *O. illyricum* is the only species with a known outcrossing rate (0.57, Michalakis *et al.* 1993). Achenes are relatively large for the thistles studied here, and as they have a small pappus are not readily wind dispersed (O'Hanlon *et al.* 1999). Due to these characteristics, they are patchily distributed with occasional long distance dispersal probably being achieved *via* animal movement, and as a fodder, wool and soil contaminant.

Achene mass & proportion of filled achenes

Variable resource access has been shown to be responsible for maintaining seed and fruit mass variation in plant species. Here, the influence of such factors has been minimised by growing plants in uniform conditions. A balanced design with 2 populations each of *O. acanthium*, *O. illyricum* and their hybrids were sampled for studying the distribution of achene mass variation. Achenes from these populations were previously collected and plants derived from them were grown outside in uniform conditions with supplemental watering at CSIRO Entomology, Canberra Australia in 10 inch pots. Flowering capitula were bagged with achenes being collected once capitula had opened after maturity. Five plants from each population were sampled with achenes collected from the primary capitulum, which is the largest capitulum growing terminal on the main stem. Twenty achenes from each were weighed to the nearest 1/1000th gram. Only those achene shells (pericarps) that contained embryos were recorded. Unfilled achenes were identified by squashing, and by having very low weights. The frequency distributions of achene masses is shown in Fig. 9.1. I examined this variation using an hierarchical ANOVA (Model II) with plants nested in populations and populations nested in species. Further analyses were performed within species using a single nesting. Hybrid sampling was expanded to four populations (2 additional sites) after the original analysis revealed that there was greater achene mass variation within the hybrid group. Coefficients of variation were calculated for achene mass variation in hybrid and parental species for pairwise comparison to test whether achene mass was equally variable in the three groups. The proportion of filled achenes was calculated by taking 10 replicates of 100 bulked achenes each for *O. acanthium* and 20 replicates of 100 bulked achenes for *O. illyricum* and counting the number of filled achenes by squashing.

Seed number

Individual *Onopordum* plants can produce a very large number achenes (each containing a single seed) and counting the number of achenes produced for each plant was not feasible. Using the same plants grown in uniform conditions, the number of seed produced per plant was estimated by determining the total achene mass produced, and dividing by the average achene mass for the individual plant. To obtain total achene mass, in addition to the bagged primary capitulum, the remaining capitula were removed from the plants and all achenes were removed and weighed. This value was corrected by determining the average mass of empty achenes and the proportion of the number of unfilled achenes relative to filled achenes within individual plants.

Proportion of germinating seed and time to germination

For germination trials, after weighing individual achenes, they were cleaned and surface sterilized in 10% household bleach, and placed in a 1.5cm deep petri dish with a filter paper substrate. To stop the moisture condensing above the filter paper on the petri dish lid, a layer of vermiculite was placed below the filter paper. The filter paper was moistened with a KNO_3 (1 mmole) and Gibberelic Acid (0.14 mmole) in solution at the first instance. Petri dish lids were snug fitting to ensure that the atmosphere around the achenes approached saturation. Filter paper was kept moistened with distilled water and achenes were left to germinate at 20°C with a fluorescent photoperiod of 8 hrs. Achenes were checked for germination by radicle emergence every 12 hrs. Twenty petri dishes for each species, each containing 40 achenes were placed in these conditions, and the number of achenes germinating in each dish was recorded.

Seedling growth rate

After germination, *Onopordum* seedlings discard their fruit coats (pericarp). As the fruit coats can comprise a significant proportion of total achene mass, fruit coats were collected and weighed to determine by difference the initial fruit and seed mass for each of 100 achenes for each taxonomic group. Correlation analyses between achene and fruit mass were performed for each of the three species groups, and pooled across groups. Time of radicle emergence was recorded, and after 21 days of growth, seedlings were dried and weighed. By difference with fruit plus seed size, the total seedling growth and growth rate were calculated. Regression analysis was performed to test whether (i) time to emergence and (ii) growth rate were determined by fruit and seed size.

Inter-specific competition during vegetative growth

Twenty-eight days after germination 180 plants were transplanted into seedling trays (substrate of sterile 1:1 sand, soil). Plants were placed in a climate controlled glass house [(temperature: average 21.3°C, maximum 25.5°C, minimum 15.2°C), (relative humidity: average 51.9%, maximum 68.9%, minimum 27.3%). Two treatments were constructed; (i) seedlings grown individually in a 5cm x 5cm x 10cm deep pot (n=30 for each group), (ii) a representative of each of the three species groups were sown together into a single pot (n=30). As different seedlings emerged on different days, only individuals that germinated on the same day were sown together. Control individuals for these treatments were established similarly. Seedlings were watered every day for the first seven days, and then every second day for the remaining seven weeks. After this time, plants were removed from the substrate, cleaned of any inorganic matter, and dried at 45°C over two days to obtain plant dry weights to the

nearest $1/100^{th}$ gram. By difference of treatment (ii) from (i) the suppression of growth due to competition could be determined and the consequences of inter-species competition could be established. To test whether each species was equally influenced by competition, two-way ANOVA was performed.

Pollen movement

To determine whether any barriers to genetic exchange through pollen movement existed, and hence to provide a potentially additional explanation for the frequency of hybrids in Europe and Australia, the movement of pollen between hybrids and parents was investigated. Healthy flowering specimens of *O. illyricum* and *O. acanthium* were harvested from the field, the day previous to the experiment. During collection, plants were placed in a box containing moistened paper to maintain the healthy condition of plants. After collection, plants were placed in a moist river-sand substrate until the following day. Plants were arranged in a circular fashion (see Figures 9.12 and 9.13) with three central plants of a known species. As individual *Onopordum* flowers are difficult to manipulate, to examine pollen movement a fluorescent dye was placed on this central plant as a pollen analogue. The dispersal of the fluorescent dye could therefore be used as an indication of pollen dispersal from the central plants to the remaining eight plants. Direct observations of insect visitation were also performed by recording the movement of an insect from one plant to the next. In this way, an overall schema of the likelihood of insects switching between species could be identified. Combined with the movement of fluorescent dye, the potential for pollen dispersal between simultaneously flowering *O. acanthium* and *O. illyricum* could be determined.

Results

Achene mass & proportion of achenes filled

O. acanthium and *O. illyricum* differed significantly in achene mass with *O. illyricum* being twice the mass of *O. acanthium* (Fig. 9.2). Hybrid *Onopordum* was intermediate between the two parental species and differed significantly from both (Fig. 9.2, Table 9.1). The distribution of variation in *O. acanthium* is narrower than *O. illyricum*, which appears to have a slightly skewed distribution (Fig. 9.1). Hybrid achenes span nearly the full range of variation for the overall group (Fig. 9.1). In addition, the distribution of achene masses in hybrid *Onopordum* resembles a bimodal distribution (Fig. 9.1). Tests for the coefficients of variation for these three groups confirmed that hybrid *Onopordum* was more variable than parental species (Table 9.2). While *O. illyricum* also appeared to be highly variable, it was not significantly more variable than *O. acanthium*. Overall, 91% of achene mass variation was partitioned among species (Table 9.1) with the bulk of achenes within species forming distinct clusters (Fig. 9.1). A significant amount of variation was also partitioned between plants within populations, with no significant differences between populations within species. When the three species were analyzed separately, most of the variation was partitioned between plants (up to 76.6% for hybrids) with no population within species differences.

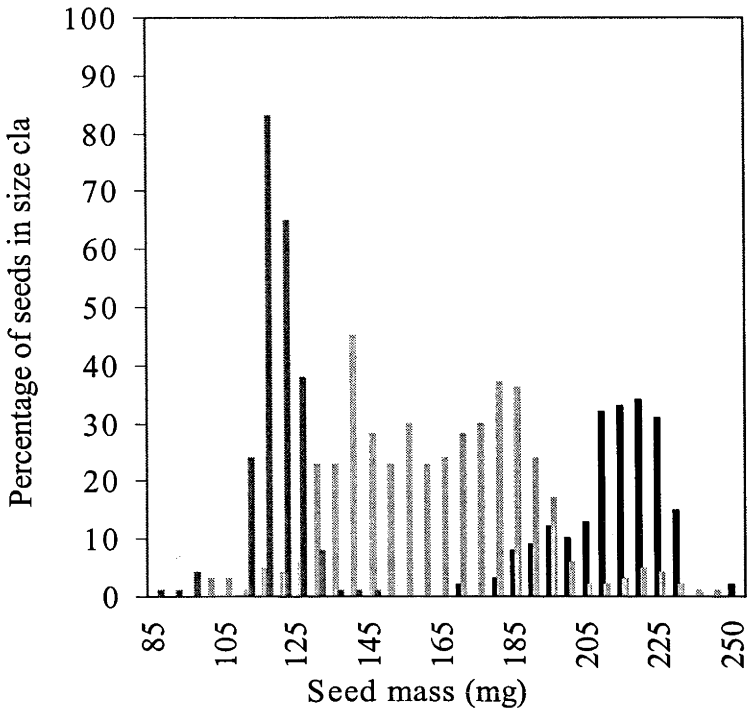


Figure 9.1 Frequency distribution of achene mass variation in hybridizing *Onopordum* (Dark gray, *O. acanthium*: light gray, hybrid *Onopordum*: black, *O. illyricum*).

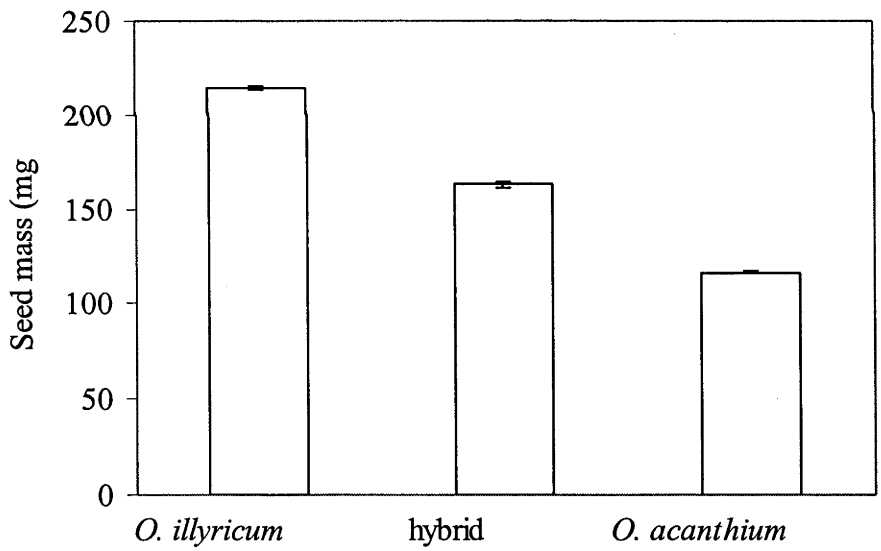


Figure 9.2 Average achene masses for *Onopordum* species and their hybrids with standard errors.

Table 9.1 Nested hierarchical analysis of variance for achene masses with 2 levels of nesting.

Source of variation	d.f.	M.S.	*F*	% Variation	*P*
Between species	2	458288	707.86	91.1	<0.001
Between pops in species	3	647	0.19	0	0.902
Between plants in pops	24	3393	60.05	6.64	<0.001
Error	570	57			
illyricum					
Between populations	1	1866.6	0.54	0	0.483
Between plants in pops	8	3457.2	43.28	67.89	<0.001
error	190	79.9			
acanthium					
Between populations	1	0.02	0	0	0.993
Between plants in pops	8	238.33	12.51	36.52	<0.001
error	190	19.05			
hybrids					
Between populations	3	5917	0.58	0	0.637
Between plants in pops	16	10227.9	6.53	76.62	<0.001
error	380	153.7			

Achene Viability	d.f.	S.S.	M.S.	*F*	*P*
Between species	2	0.713	0.3566	22.93	<0.001
Error	37	0.5755	0.01555		

O. acanthium (94%) and *O. illyricum* (92%) did not differ in the proportion of filled achenes. However, hybrid *Onopordum* (73%) had approximately 20% fewer achenes filled (one way Kruskal-Wallis analysis of variance on ranks, d.f=2, $p<0.01$). Combined with achenes mass variation, individual hybrid plants were far more variable in their size than their parent plants.

Table 9.2 Tests for equality of coefficient of variation for achene mass and proportion of achene filled with pairwise tests.

Comparison	d.f.	X^2	*p*
Species achene mass	2	19.743	<0.001
illyricum v hybrid	1	5.561	0.025
acanthium v hybrid	1	13.719	*<0.001
acanthium v illyricum	1	4.844	0.050

*Significant after Bonferonni correction.

Achene number and production

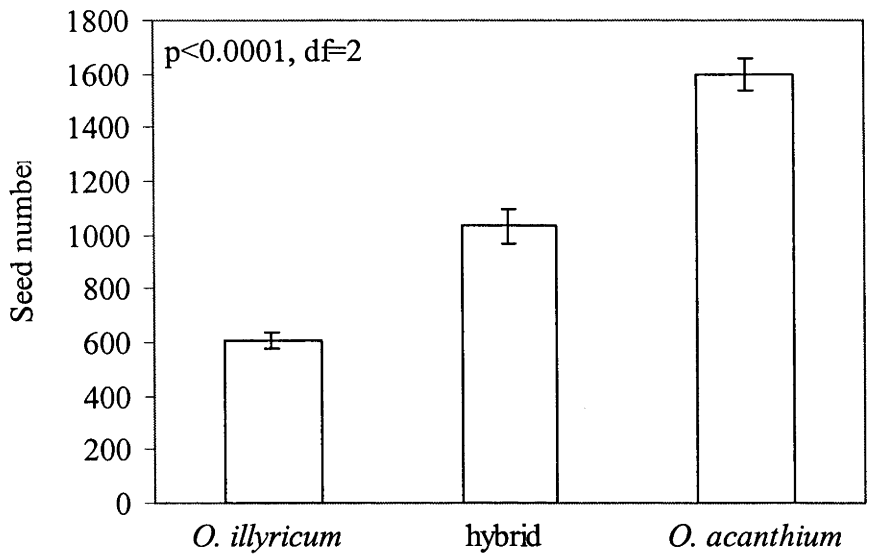


Figure 9.3 Average number of achenes produced in a controlled environment for two parental *Onopordum* species, and later generation hybrids with standard errors (results of one way Kruskal-Wallis analysis of variance on ranks given).

O. acanthium produced nearly two times the achenes produced by hybrids and nearly three times that by *O. illyricum* (Fig. 9.3). In a natural context it is likely that the between plant variation is more significant than represented here due to fine-scale variation in ecological conditions. When reproductive output was measured as total achene mass produced (Fig. 9.4), there was less disparity between the groups, with *O. acanthium* and hybrids not being significantly different from each other (Dunn's pairwise comparison, $p<0.001$)

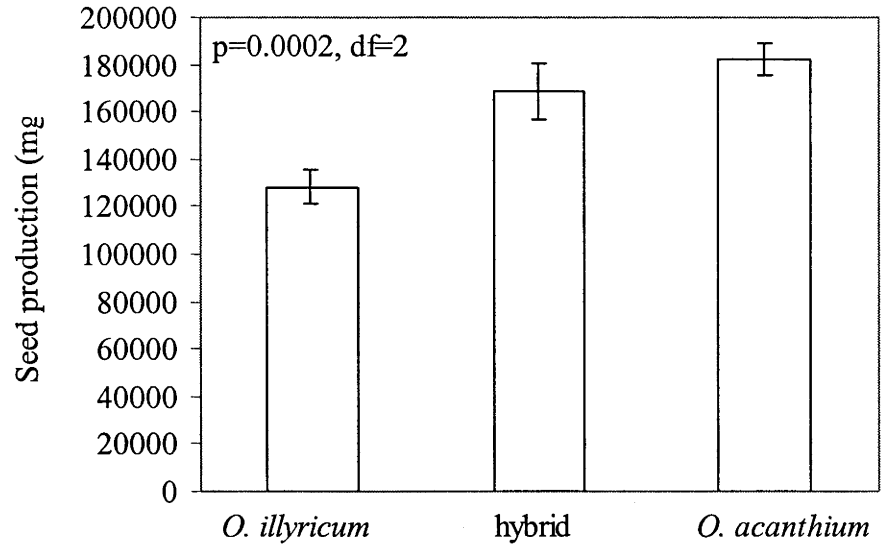


Figure 9.4. Average total mass of achenes produced per plant for two parental *Onopordum* species, and later generation hybrids with standard errors (results of one way Kruskal-Wallis analysis of variance on ranks given).

Proportion of germinating achenes and time to germination

O. acanthium and *O. illyricum* had germination rates of over 90%. In contrast, hybrids had a significantly lower germination rate of only 74% (Fig. 9.5).

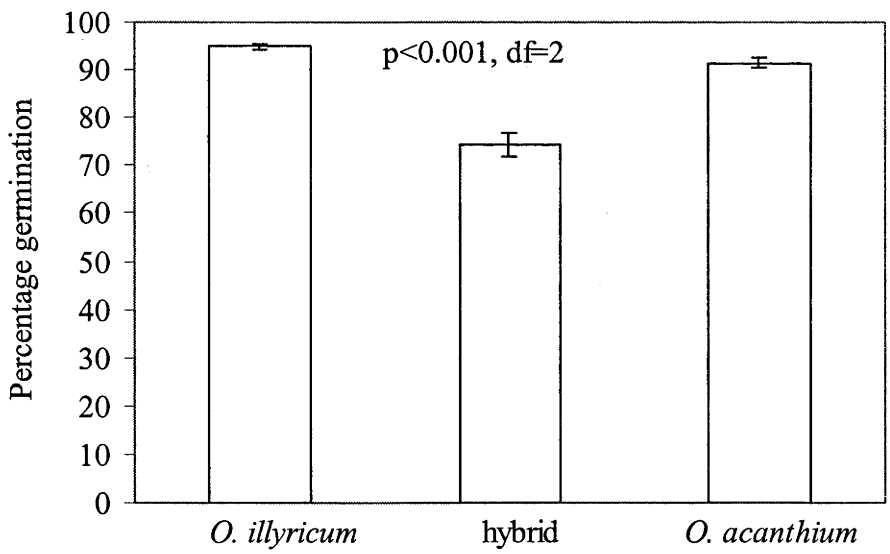


Figure 9.5 Percentage germination of achenes for two parental *Onopordum* species, and later generation hybrids with standard errors (results of one way Kruskal-Wallis analysis of variance on ranks given).

The rate of germination of achenes from different groups followed a very similar response curve, with the only difference between the groups being the lag before

germination commenced (Fig. 9.6) The groups with the larger achenes, which also have the larger fruit and the larger fruit coats had slower germination. Radicle emergence began on day three for *O. acanthium*, day four for hybrids and day five for *O. illyricum* with germination complete by day seven for *O. acanthium*, ten each for hybrids and for *O. illyricum*.

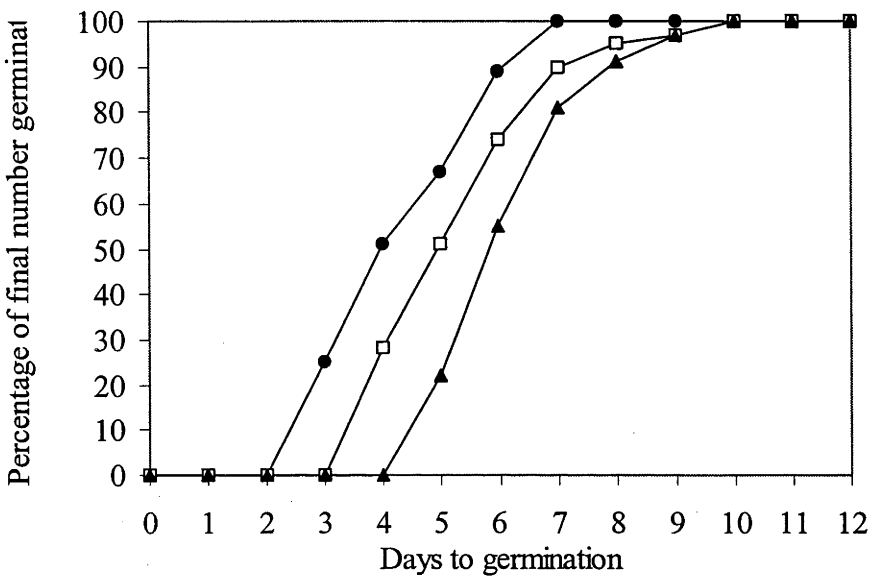


Figure 9.6 Germination curves for two *Onopordum* species and later generation hybrids.

A positive correlation between fruit and entire achene mass was identified when all data was pooled (coefficient 0.435, R^2 0.844, $p<0.001$). A similar relationship was identified for individual species (*O. acanthium* coefficient 0.399, R^2 0.714, $p<0.001$; *O. illyricum* coefficient 0.537, R^2 0.381, $p<0.001$; hybrids coefficient 0.482, R^2 0.463, $p<0.001$). Further regression analyses were conducted using both entire achene mass and fruit plus seed mass. Similar patterns were obtained from both, with stronger relationships for fruit and seed size. Consequently results only from fruit and seed mass are provided.

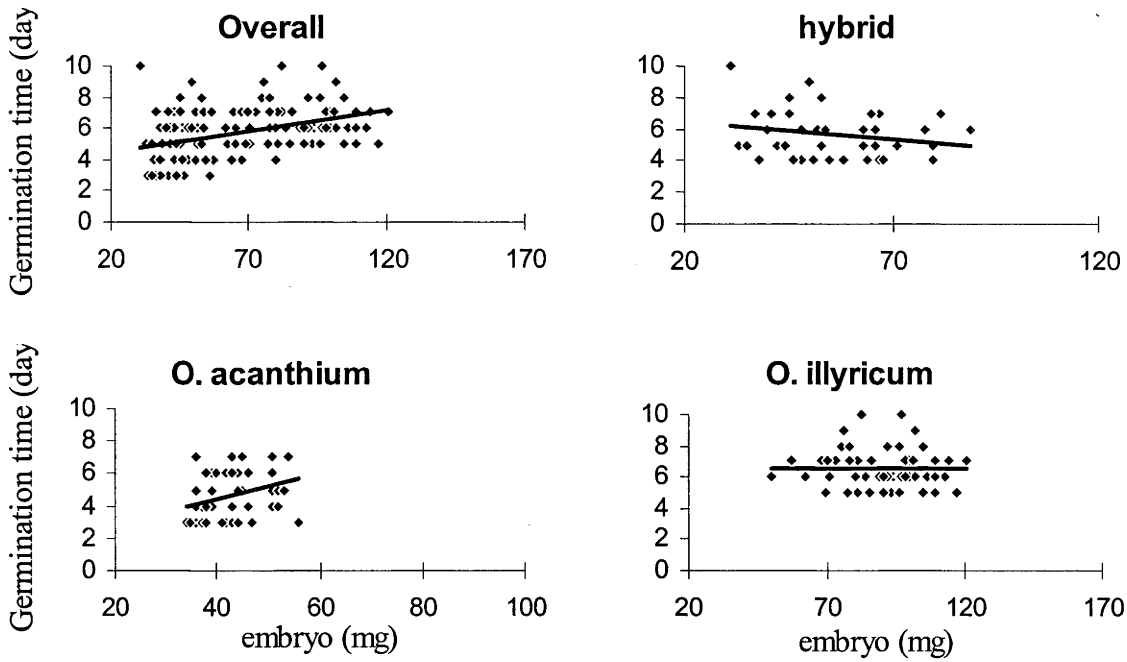


Figure 9.7 Regression analyses of germination time against fruit size for (i) pooled data y=0.027x+3.843, R^2 0.180, p<0.001; (ii) hybrids y=-0.020x+6.763, R^2 0.038, p=0.236; (iii) *O. acanthium* y=0.075x+1.441, R^2 0.103, p=0.010; (iv) *O. illyricum* y=-0.001x+6.610, R^2 0.001, p=0.93047.

Overall, germination time was positively related with fruit and seed mass with substantial variance about the regression, reflecting the different times to germination for the different species (Fig. 9.7), each of which have different average fruit-seed masses. Within species groups there was no overall trend to time to germination with embryo masses. For *O. acanthium* there was a significant positive relationship (p=0.01). However, for hybrids there was a negative trend, and for *O. illyricum* there was no relationship.

Seedling size and growth rate

While different species had different seedling sizes after 21 days (Fig. 9.8), this was largely due to different fruit-seed sizes for the groups because after 21 days, seedling size was not very different from fruit-seed size, the mass of which formed the bulk of the seedling. When growth was measured as a percentage of fruit-seed size rather

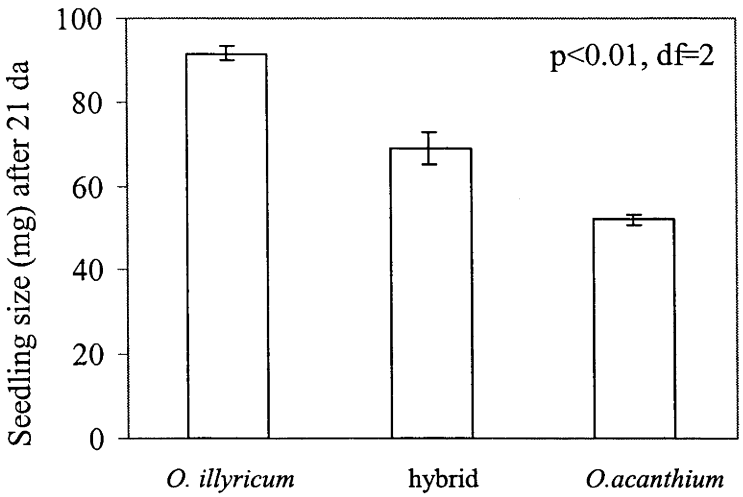


Figure 9.8 Average seedling size for two parental *Onopordum* species, and later generation hybrids with standard errors (results of one way Kruskal-Wallis analysis of variance on ranks given).

than seedling size, there was no general relationship between growth and fruit-seed size (Fig. 9.9). While a significant negative relationship between fruit-seed size and growth rates was observed when species were pooled, this result merely reflected differences in the averages of species growth rates in these conditions, particularly the low rate of growth of *O. illyricum* (Fig. 9.10), which itself showed a strong negative correlation between fruit-seed size and growth rate (Fig. 9.9).

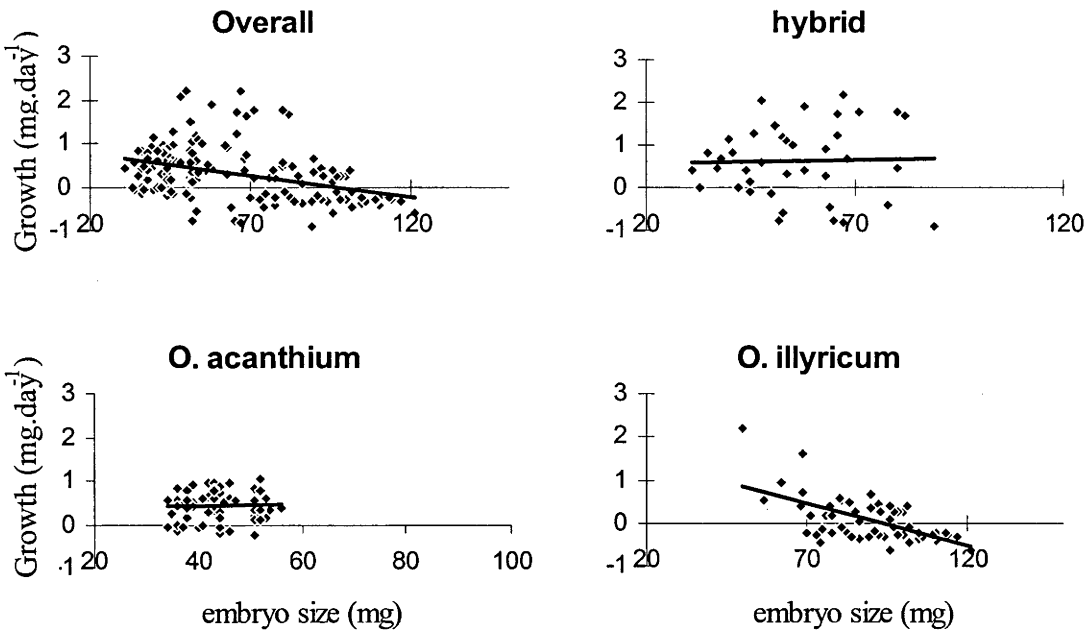


Figure 9.9 Regression analyses of growths rates on fruit-seed sizes for (i) pooled data y=-0.010x+0.952, R^2 0.157, p<0.001; (ii) hybrids y=0.002x+0.519, R^2 0.001, p=0.861; (iii) *O. acanthium* y=0.002x+0.351, R^2 0.001, p=0.821; (iv) *O. illyricum* y=-0.020x+1.860, R^2=0.391, p<0.001.

Under the conditions used in this study, hybrids were found to grow at a significantly higher rate than either of the parental species, with *O. illyricum* showing very poor growth (Fig. 9.10). Such poor growth by *O. illyricum* probably did not reflect any intrinsic low fitness but probably a limitation of the study, with unsuitable conditions for growth in *O. illyricum*. This is supported by the strong negative relationship between growth rates and embryo sizes for *O. illyricum*, with larger fruit-seed, showing negative growth. This indicates that there may have been insufficient resources available to sustain larger fruits, preventing any growth and inducing senescence.

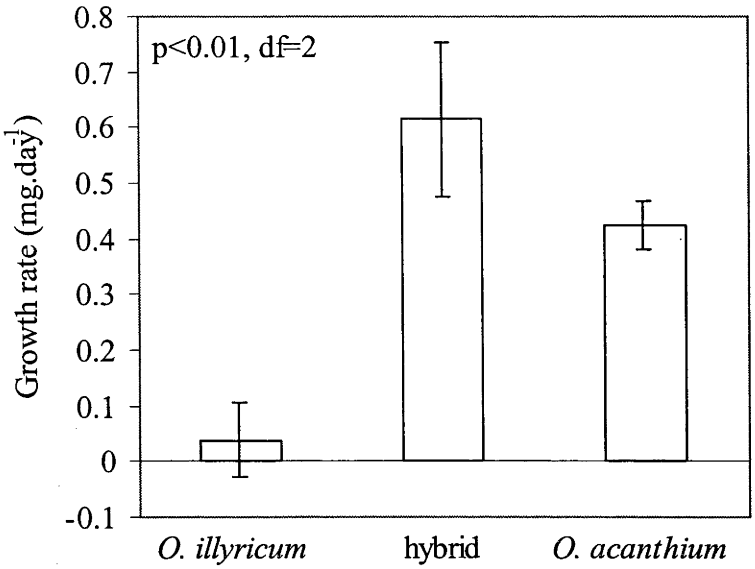


Figure 9.10 Average growth rates for two parental *Onopordum* species, and later generation hybrids with standard errors (results of one way Kruskal-Wallis analysis of variance on ranks given).

Inter-specific competition during vegetative growth

Inter-specific competition was confirmed by the decrease in mass of plants grown together when compared with growth in the absence of competition (Fig. 9.11, Table 9.3). When grown in the absence of competition, there was a significant difference between species, with *O. acanthium* being smaller than either hybrids or *O. illyricum*, which showed no difference (Fig. 9.11, Table 9.3). This result contrasts to the seedling size after 21 days, and reflects the improved conditions for growth. Competition induced a decrease in the growth of all plants of equal proportion (Fig. 9.11) and analysis of variance revealed no interaction between species and competition with all species affected equally by competition (Table 9.3). However, for total plant size in competition, hybrids and *O. illyricum* achieved the greatest size.

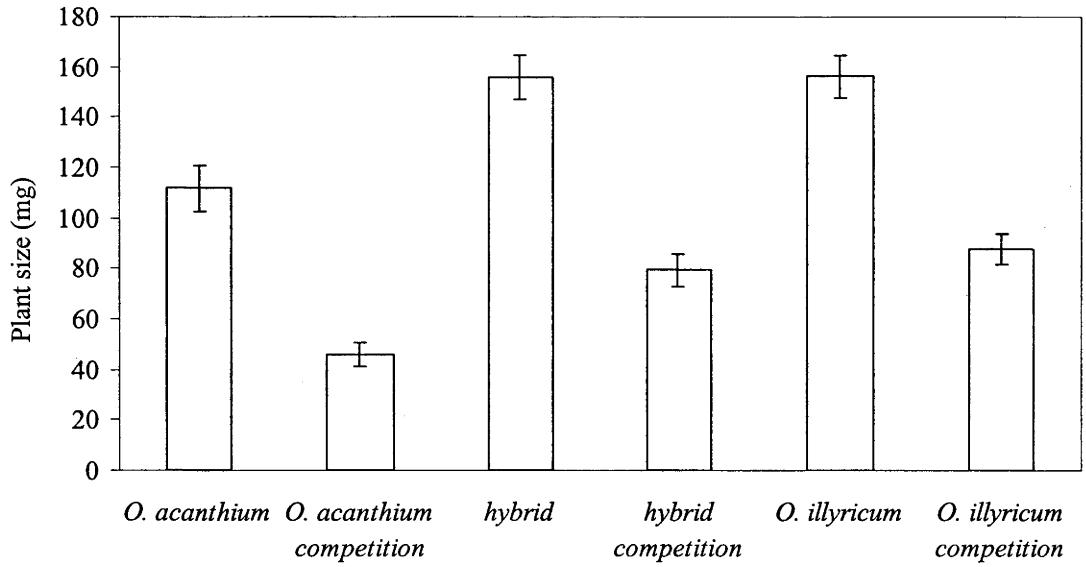


Figure 9.11 Average plant size for two parental *Onopordum* species and later generation hybrids for competitive and non-competitive regimes with and standard errors.

Table 9.3 Two-way analysis of variance for plant sizes between different species with and without competition.

Source of Variation	*SS*	*df*	*MS*	*F*	*p-value*
Competition	208046.1	1	208046.1	136.5475	<0.001
Species	63225.57	2	31612.79	20.74851	<0.001
Interaction	889.4762	2	444.7381	0.291896	0.747
Within	246826	162	1523.617		
Total	518987.1	167			

Pollen movement

Some variation in the numbers of fluorescent particles detected on 'sink' plants was identified ranging between zero and six grains. Overall, these numbers were very low and I anticpated that as I placed probably greater than several hundred grains on the source plants that more particles would be detected. Insects were active throughout the recording period indicating that the fluorescent dye was either poorly picked up by insects from source plants, transfer was low between plants, or that grains did not adhere to the stigmas of other floral parts of the sink plants. Nevertheless, there was no apparent difference in the movement of fluorescent dye between different species.

Figure 9.14 shows that particle movement was not significantly different irrespective of the source/sink species combinations.

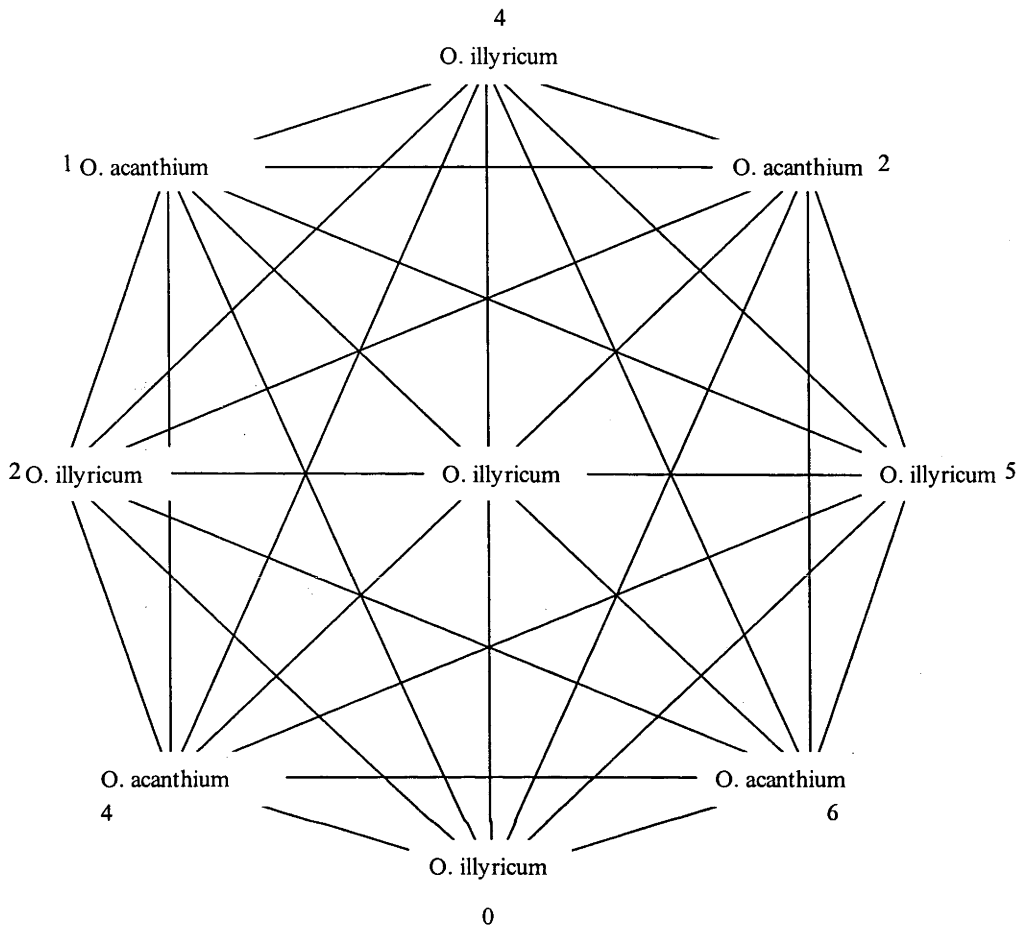


Figure 9.12 Arrangement of plants in a circle surrounding three 'source' *O. illyricum* plants. Source plants have a fluorescent dye which can act as a pollen analogue. Lines connecting plants show potential pollen movement, and numbers represent the number of fluorescent particles detected on individual plants.

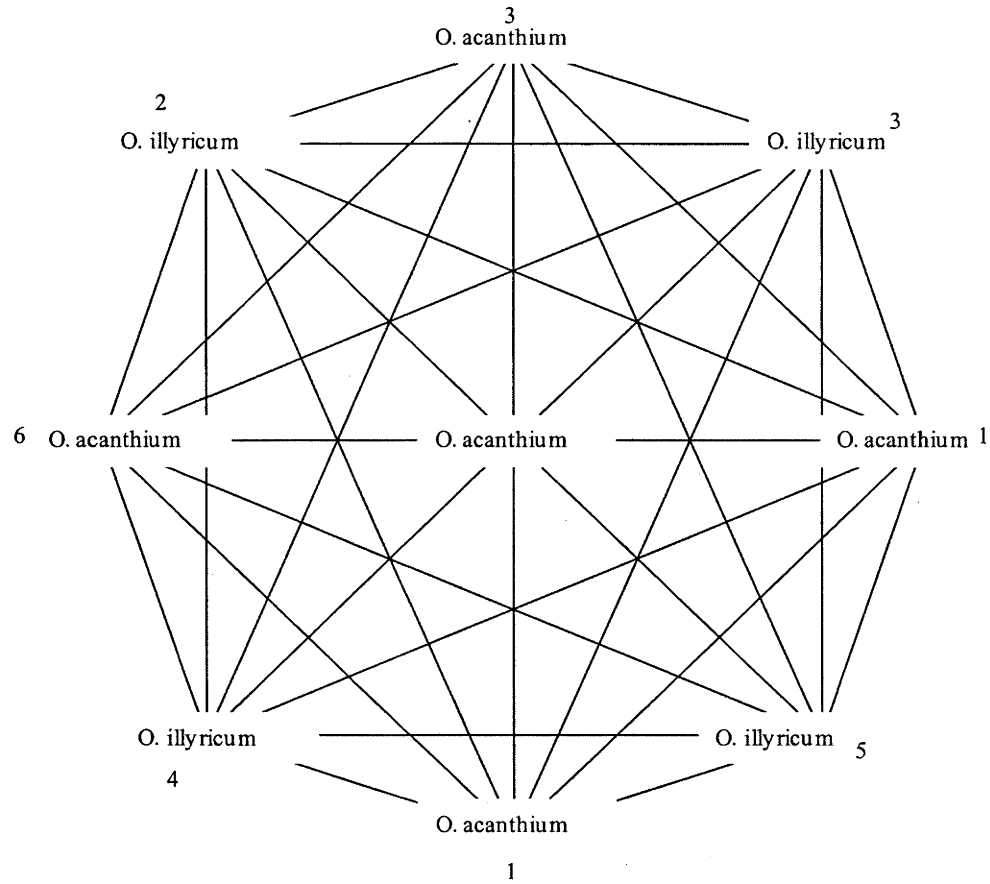


Figure 9.13 Arrangement of plants in a circle surrounding three 'source' *O. acanthium* plants. Source plants have a fluorescent dye which can act as a pollen analogue. Lines connecting plants show potential pollen movement, and numbers represent the number of fluorescent particles detected on individual plants.

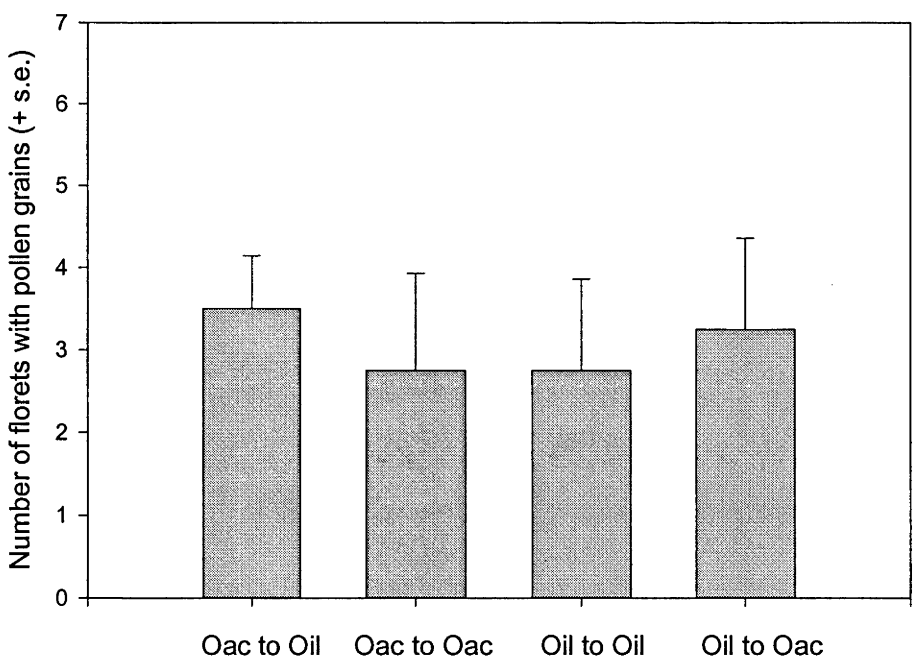


Figure 9.14 Average number of grains of fluorescent dye detected on possible 'sink' plants after depositing dye on 'source' plants (Oac – *O. acanthium*, Oil – *O. illyricum*) with standard errors.

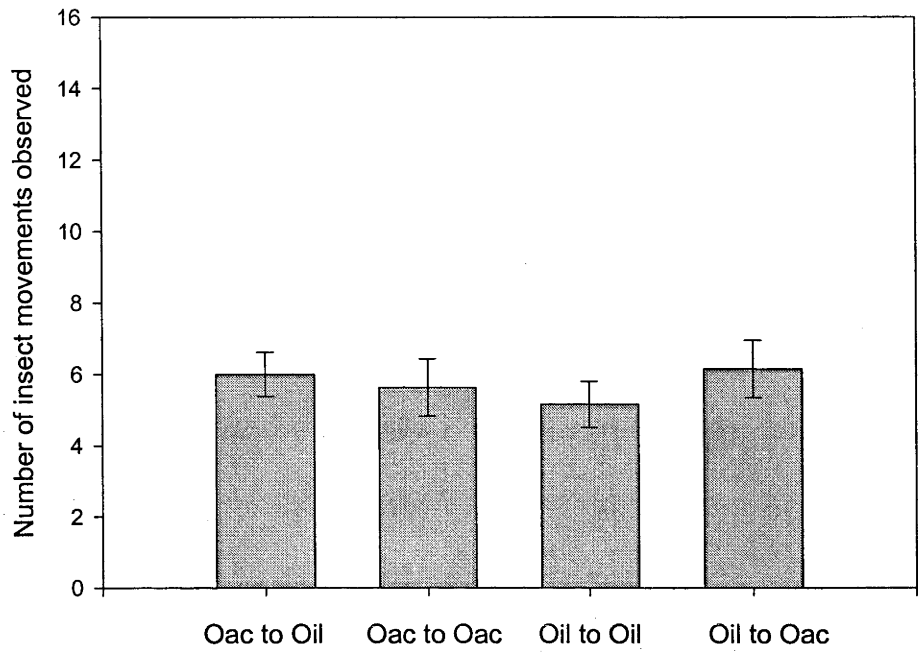


Figure 9.15 Average number of insects moving from 'source' plants to 'sink' plants (Oac – *O. acanthium*, Oil – *O. illyricum*) with standard errors as an indication of gene flow between species.

Analysis of the movement of fluorescent pollen analogue particles was supplemented by direct observations of insect visitations (Figure 9.15). Records of movements from one plant to another show no preferential dispersal within particular species, and insects were just as likely to move from one *Onopordum* species to another *Onopordum* species.

Discussion

Achene mass variation between species

Onopordum hybrids have a novel, intermediate achene mass (Table 9.1, Figs. 9.1 & 9.2). As plants were grown in uniform conditions, there is a clear genetic basis for achene mass in this hybrid complex, with parental species and their hybrids all

showing significant differences from one another (Table 9.1, Fig. 9.2). Apart from achene mass differences between species, within species there were also significant differences in the average achene masses from different mothers (Table 9.1). Unlike many species where within-plant variation is prominent (Mendez 1997, Vaughton & Ramsey 1998) most of the within species variation was partitioned between individual plants within populations (Table 9.1). As resource patchiness was not responsible for achene mass differences, such a pattern of variation suggests that the maternal genotype was critical for determining achene mass in *Onopordum*. Additionally the variation within hybrids was greater than that observed within parental species (Figure 9.1, Table 9.1 & 9.2), probably reflecting the segregation of alleles from different parents associated with achene mass. The increased variation in hybrids may also reflect relatively recent hybridization and insufficient time for selection to eliminate either extremely large or small achenes, each of which may be associated with lower fitness. Alternatively, selection may be destabilizing in hybrid populations, selecting for parental achene masses. There is some evidence that achenes from hybrid plants studied here have tended to resemble one parent more than the other with an apparently bimodal distribution (Fig. 9.1). This could be taken as evidence of selection favoring parental achene masses. However, it is common for later generation hybrids to display a morphological affinity with one parent over the other through the segregation of alleles and non-additive character expression (Rieseberg 1999).

The distribution of achene masses (Fig. 9.1) could equally represent the same scenario. The complexities of genetic rearrangements or interactions in hybrids may also revealed by the increase in the proportion of unfilled achenes and lower viability. These may be produced by genetic incompatibilities within zygotes, or between

maternal and offspring genotypes. However, without knowing more about the patterns of pollination and specific breeding system information, it is not possible to determine whether genetic incompatibilities or differential abortion and allocation of resources was responsible for such an effect. However, preliminary studies on pollen movement show that there are no barriers to pollination between these two species, and insects capable of pollination do not appear to discriminate between parental species (Figures 9.14 and 9.15). Finally, variation in offspring mass may be adaptive in hybrid derivatives. As there is a range in plant genotypes and probably also phenotypes, and plant genotype strongly influences achene mass, there may be an advantage in maintaining achene mass variation. Other factors such as different dispersal mechanisms, or germination times (Venable & Brown 1988, Geritz 1995), by interactions between plant size and reproductive behaviour leading to assortative mating (Venable 1992, Sakai & Sakai 1995), interactions with herbivores (Moegenburg 1996), through preferential allocation of parental resources to high-quality genotypes (Temme 1986, Waser 1993) or hormonal control of the parent by the developing zygote (Lipow & Wyatt 1999) not studied here may also contribute significantly to plant fitness. Therefore it is not possible to determine in this study the ultimate causes and consequences of increased achene mass variation in hybrids.

Consequences of achene mass variation to fitness in hybridizing genotypes

The two parental taxa naturally grow in different regions (Franco 1976) with marked climatic differences, and it is likely that they have different optimal achene masses appropriate for their respective habitats. Hybrids have a generally intermediate though more variable achene mass. Hodgkinson *et al.* (1998) showed a general trend between

achene mass and habitat shade, as well as a negative correlation between seed mass and longevity. While specific studies have not been performed for the *Onopordum* species studied here, it is likely that similar differences may also occur. The differences in the abundance of hybrids and parental species in different continents suggests a role for environmental variation in determining hybrid fitness (Hatfield & Schluster, 1999). Hybrids may therefore have a unique optimal achene mass as a consequence of having a different habitat preference, as well as a different phenotype. An alternative explanation may be that instead of selection favoring an intermediate hybrid achene mass, there may be relatively relaxed selection for achene mass in *Onopordum*. This is supported somewhat by the lack of relationships between fruit-seed mass and various fitness components, but if larger achenes are no fitter than smaller achenes, selection should favor plants producing a greater number of smaller achenes (Smith & Fretwell, 1974). Furthermore, the overall correlation between fruit-seed size and seedling size, as well as the decrease in the number of achenes produced in plants with larger achenes indicates that a traditional trade-off between selection for achene mass and number might occur in *Onopordum*. That the two parental species differ markedly in achene mass and are capable of hybridization make it likely that achene mass has been constrained by selection.

Regardless of the causes of the novel achene mass in hybrids, *Onopordum* acts as a demonstration that an altered offspring mass may be introduced rapidly through hybridization between once separated lineages, and not necessarily through optimizing selection for the balance between offspring mass and offspring number. Where hybridization occurs between species or populations with different offspring masses, a larger spectrum of offspring masses may be possible, depending on the

inheritance and outcome of hybridization. It is commonly observed that species displaying seed mass variation show differentiation between populations (Zhang & Hamill 1997, Zhang 1998, Telenius & Torstensson 1999) and consequently, such hybridization between populations may be an important factor in the introduction of seed mass variation.

Clearly there is a genetic basis, expressed primarily between maternal genotypes with such variation maintained for many generations. In the context of plant invasion, the increase of intra-specific variation may contribute to the rate of adaptation to new conditions. Such invasion may be either geographic range expansion or persistence in a habitat that becomes modified. Susko and Lovett-Doust (2000) demonstrated extreme variation within weedy garlic mustard (*Alliaria petiolata*). It was suggested that such variation was important to survival during recruitment with a trade-off between emergence time and growth (leaf production and plant height). When variation is present, a subsequent change in selection regime may favor different offspring masses. In the case of hybridization in *Onopordum*, such an altered regime may select for an achene mass that did not exist prior to hybridization. As hybridization can rapidly introduce novel variants, hybridization may contribute significantly to adaptation in invasive plants.

Offspring number, viability, germination time, growth rate and hybrid fitness

Traditionally, life history traits such as germination, growth and reproductive output have been measured to infer fitness of hybrids relative to parents. Such studies have usually compared genealogical classes such as F_1, F_2 and backcross genotypes with parental genotypes (see Arnold & Hodges 1995). However there is now considerable

data illustrating that later generation hybrids can have different relative fitness to such early generation hybrids (Arnold & Hodges 1995, Rieseberg 1999). Provided that later generation hybrids can become established, such as in those studied here, it is the relative fitness of these later generation hybrids that will determine the evolutionary significance of hybridization. While ideally hybrids with known genealogy should be studied, their production can be a very time consuming process, over many generations. Here an alternative approach of sampling achenes from the field from several populations from later generation hybrids was employed. Such a scheme, while not being capable of differentiating various effects of specific mating types, can give a realistic interpretation of long-term consequences of hybridization for fitness in the described traits.

For achene production in a controlled setting, *O. acanthium* plants produced nearly twice the number of achenes than did the hybrid plants, and nearly three times that of *O. illyricum* (Fig. 9.3). Such a result may reflect the different number of capitula produced per plant, a higher capacity to acquire resources or to convert resources into offspring, different energy allocations into flowering or architectural structures, or simply different achene sizes with similar energy allocations. Regardless, hybrids produce an intermediate number of achenes, and are neither intrinsically nor uniformly unfit in this respect. While hybrids produced a higher proportion of unfilled achenes, probably as a consequence of a genetic incompatibility, they still produced more achenes than *O. illyricum*. When total reproductive output was considered, hybrids were not significantly different from *O. acanthium* but significantly more productive than *O. illyricum* (Fig. 9.4). With germination, hybrid achenes showed the lowest fitness and were significantly lower than *O. acanthium* (Fig. 9.5), and hence in

this trait hybrid unfitness is most pronounced. However, in a natural setting, this may not relate directly to fitness, as mortality in *Onopordum* is greatest during establishment from germinated seeds (Groves *et al.* 1990, Pettit *et al.* 1996). While not significantly different from *O. illyricum*, in the field this reduction in vigour may result in strong selection against the bulk of hybrid *Onopordum*. Despite this potential, a proportion of hybrid achenes showed rapid germination and since there is such a high mortality at germination, it is still likely that early germinating hybrid seed may survive. Additional trials in the field, preferably using F_1 hybrids would be required to test the overall importance of such a reduction of fitness in this trait. Furthermore, in a natural setting, the variation in achene number between taxa may be less significant due to interactions between plant size and local conditions.

One fitness trait that has previously been shown to relate to mortality is plant size, with evidence that large plants of *O. illyricum* were fitter than smaller plants (Rees *et al.* 1999). Therefore, aspects relating to plant size such as embryo size, growth rates, and also possibly competitive ability probably contribute to plant fitness. Fruit-seed mass for hybrids again was intermediate between parental species, and in petri dishes placed in a germination cabinet, hybrids showed a significantly higher growth rate than either parental species for the first 21 days (Fig. 9.8). However, after this time *O. illyricum* still had the largest seedlings despite showing very low growth rates. After transplantation into soil in a glass house, and growth over an eight week period, hybrid plants were as large as *O. illyricum* and significantly larger than rosettes of *O. acanthium* (Fig. 9.11). In competition, hybrid plants showed a similar decrease in plant mass as parental species, with each species showing the same proportionate decrease in growth (Fig. 9.11). By virtue of embryo size, growth rates, competitive

ability and final plant size, hybrid *Onopordum* was not intrinsically less fit than parental species.

Without further studies of the importance of the various traits to overall long-term fitness, and assigning weights to characters, it is not possible to determine whether a multiplicative estimate for hybrid fitness (such as that used by Hauser *et al.* 1998) would give an accurate estimate of fitness for the present study. In addition, any trait weightings assigned may be influenced by environmental variation, and it is likely that such overall estimates would change depending on the environmental conditions employed. Nevertheless, as an overall summary of fitness, in most traits hybrids usually displayed equal or higher fitness than the parental species in the given conditions. Importantly, for most of the traits examined, hybrids also displayed a greater degree of variation than the parental species (cf. Figs. 9.2, 9.3, 9.4, 9.5, 9.8, 9.10 & 9.11). Such increased variance may allow for a greater degree and rate of adaptation in hybrid *Onopordum* to new environmental conditions.

This study has raised several important similarities with studies in other plant and animal species. A difference between the relative fitness of stickleback hybrids (fish) in controlled and in natural conditions was used to infer environment-dependent selection (Hatfield & Schluter 1999) for hybrid fitness. When grown in uniform conditions, hybrids performed relatively well. However, when grown in natural conditions, hybrids were out-competed by parental stickleback species. Such a difference was taken as evidence that there was no intrinsic unfitness in hybrids, but that ecological conditions acted to promote species boundaries in nature. A similar argument could be applied for hybrid *Onopordum*. In the controlled conditions employed here, hybrid *Onopordum* performed relatively well, showing decreased

fitness in only the proportion of achenes filled and in achene viability. However, in the native European range, hybrids occur less frequently than parents and only in the presence of parental species and form a distinct and narrow hybrid zone at the boundary between the species (see Gonzalez-Sierra 1992, *Chapter Eight*). If hybrids were equally fit in Europe, they may be expected to merge with parental taxa, leading to a breakdown in species boundaries and the formation of a hybrid swarm. In contrast to Europe, hybrids are abundant in Australia, and occur frequently in the absence of parental species (O'Hanlon *et al.* 1999, see also *Chapters Six* and *Seven*) indicating at least a release from selection pressures against hybrids. Furthermore, there is some evidence that in Australia, hybrids are fitter as selection favoring heterozygotes over time within a cohort has been documented in a hybrid population (Groves *et al.* 1990). Therefore, given their moderately high intrinsic fitness, the apparent differences between the relative abundance of hybrids between continents could be taken as evidence that the European ecological context has acted to separate the two parental species, but in Australia such conditions may not exist. Such a pattern is consistent with a bounded hybrid superiority model, which may be most evident in invading species. Hybrids between a native *Carprobrotus chilensis* (L.) and an alien *C. edulis* (L. Bolus) showed either an increase in, or similar fitnesses to parents in fruit weight, seed set, achene mass, germination rate, survival, growth and root to shoot ratios (Vila & Dantonio 1998), a similar range of traits as those studied here in *Onopordum*. Indeed hybridization between natives and aliens has been cited as a potentially important cause of extinction (Levin *et al.* 1996). Such studies point to the importance of further studies of the influence of invasion and environmental variation on hybrid fitness in *Onopordum*.

Chapter Ten *Conclusions*

In concluding this thesis, I endeavor to integrate the preceding research chapters and to highlight several aspects of the evolution and ecology of *Onopordum* thistles with particular emphasis on hybridization and invasion. Such a task requires the summary of the key findings of each chapter, with the exception of *Chapters Two and Three* which provide technical details of molecular tools and their application to weed research in general.

In *Chapter Four*, I conducted a phylogenetic analysis of a subset of the Carduinae thistles, and attempted to place the genus *Onopordum* within an evolutionary tree. Not only did this provide a suitable avenue for exploring the utility of several genetic markers for the research planned, it also provided opportunities for achieving new insights into the phylogeny of this group. In terms of the development of genetic markers, this work was extremely important in that it led to the development of a new molecular test for identifying size homoplasy among fragments derived from AFLP (and possibly other multilocus fingerprinting technqiues). In addition, two usually variable cpDNA markers were sequenced, revealing high levels of variation, of which only a low proportion was phylogenetically informative. Taken together with morphological data, the overall resultant phylogeny was neither satisfactorily resolved or robust. Such a result may be the consequence of AFLP markers and cpDNA evolving at inappropriate rates for effective reconstruction, but it may also reflect non-hierarchical evolution in the group. The latter hypothesis was supported by the fact that many informative molecular characters were identified, but these gave conflicting evolutionary signals. Despite the overall phylogeny being equivocal, several minor clades within it were revealed as relatively robust. At an applied level, the resultant

phylogeny may inform tests for host-specificity testing in this important group of thistles. Such testing procedures should reflect the phylogenetic complexity of this group (Wapshere 1974).

It was hoped that a study of the Carduinae thistles (*Chapter Four*) would provide informative cpDNA markers with which to explore the phylogeny and phylogeography of the *Onopordum* genus of thistles that have colonised Australia. Ideally cpDNA and nuclear data would have been used in concert to determine the maternal and paternal evolutionary patterns of this group but no cpDNA variation was detected within *Onopordum*. Such a finding is of itself interesting. *Asteraceae* is a modern family and genera within it may therefore be relatively uniform. However, given the broad range of morphological diversity in *Onopordum*, such a result was unexpected, and may point to either very rapid morphological evolution, or to cpDNA homogenization through selective sweeps and reticulation. Recently, techniques for sampling more variable cpDNA sequence have been identified providing an interesting avenue for testing these hypotheses.

Despite the lack of cpDNA variation, AFLP fingerprinting produced many variable markers, and using the test for homoplasy described in *Chapter Three*, it was shown that phylogenetic reconstruction would not be confounded by size homoplasy. The number of markers enabled the identification of several samples previously identified as pure parental stock as being of hybrid origin. Such a result potentially points to more extensive hybridization between species of *Onopordum* in Europe than previously expected.

Phylogenetic reconstruction of AFLP data revealed three major clades that correponded with geographic regions; temperate Europe, the Iberian Peninsula, and the Aegean/Balkan region. These latter two areas have been shown to be important refugial areas during glacial maxima, suggesting an evolutionary pattern structured by such climatic oscillations. It is possible that other more ancient processes may account or influence the current architecture, but the correlation with geographic patterns in the last glacial maximum provides a economical and satisfying explanation. Such a pattern has implications for the distribution of morphological characters in the group. While genetic data shows that species within regions are more closely related to each other than to species in other regions, morphological data would contradict this finding. Reconciliation between data sets may be provided by the explanation of lineage sorting and recolonization. Such a pattern could emerge where widepsread hybridization leads to morphologically diverse refugial populations, and subsequent recolonization of Europe with selection by a range of conditions sorting such variation, leading to convergence.

The work outlined in *Chapter Five* not only suggests that hybridization has been important in the evolution of the group, particularly during invasion, but also contributes to the study of the evolution of host-choice by invertebrate herbivores within the genus. A similar biogeographic pattern in the composition of the insect fauna has been observed, with the fauna found on species from the same region being more similar to that found between groups of apparently related but isolated species (Briese *et al.* 1994). Also, speciation within a group of weevils of the genus *Larinus* has been shown to have been driven by biogeographic isolation between the Iberian and Aegean regions rather than host-plant specialization within the genus (Briese *et*

al. 1996). The identification of similar patterns between the flora and fauna is particularly satisfying in terms of evolutionary studies and points to coevolution.

The identification of several diagnostic markers for each European species of *Onopordum* (*Chapter Five)* proved to be very useful in determining the identity of taxonomically ambigous, intermediate samples from Australia. In *Chapter Six*, it was shown that most Australian populations of *Onopordum* contained individuals displaying segregation of AFLP markers, diagnostic primarily for *O. acanthium* and *O. illyricum*. Such a finding is strong evidence that the putative hybrids identified in Australia and described as *O. illyricum* sp. aff. are indeed derived from matings between *O. acanthium* and *O. illyricum*. However, intermediate individuals also contained several fragments that were identified only in species other than *O. acanthium* and *O. illyricum* sampled from Europe. This may indicate that multi-species introgression has occurred in this group, and opportunities for such introgression would be high in certain contexts (especially cultivation) prior to invasion. Despite this possibility, sampling of European material was restricted, and it could not be ruled out that that some genetic markers apparently diagnostic for species other than *O. acanthium* and *O. illyricum* may also be found at low frequency in these two species. Finally, Australian hybrids also contained fragments absent from any collections of European *Onopordum*. This may either reflect post-hybridization genetic rearrangements, or simply the limited sampling within Europe. It is unclear whether such an evolutionary history characterised by hybridization has any special implications for biological control in Australia. However, differences between the parental species (such as capitulum size, longevity of rosettes, achene mass and

number and leaf pubesence) may all contribute to hybrids having different responses to different management strategies (O'Hanlon *et al.* 2000)[b].

More detailed investigations of population genetic structure showed that there was relatively little genetic diversity present within populations, with most populations being strongly differentiated from each other (*Chapter Seven*). Even for many populations containing hybrids, I identified little within-population genetic variation when compared with that found between populations. Finally, populations containing hybrids were found to be comprised exclusively of hybridized genotypes. Such patterns of population structure indicate that populations are generally established by one or a few similar hybrid genotypes. Had parental species come into contact at the same location, one may have expected more genetic diversity within populations. As *Onopordum* populations may be capable of very rapid increase in numbers of plants (by virtue of their high achene set each year), the later migration of one or few seed into an already established population may not contribute significantly to allele frequency change (unless selection was very strong).

To complement this research, in *Chapter Seven* I also showed by spatial autocorrelation analysis of morphological variation that most populations were similar in identity to those neighboring it, with decreasing similarity up to a distance of 60km. Such autocorrelation provides additional evidence that many sites have probably been colonized by local gene flow, and that hybrid sites may have had founders from a neighboring hybrid population. Such a finding provides support to the hypothesis that limited variation in hybrid sites is a consequence of having few hybrid founders. *Onopordum* therefore provides an example of where the invasion of new regions by

hybridizing plants can lead to the formation of a suite of distinct and practically isolated lineages, suggesting that hybrid speciation need not always be a special case of sympatric speciation (except under the biological or recognition species concepts).

The identification of a previously unknown hybrid zone between *O. acanthium* and *O. illyricum* in southern France has important implications for the evolutionary consequences of hybrids in this group (*Chapter Eight*). Narrow hybrid zones usually indicate that strong barriers to gene flow and introgression exist between species. The narrow hybrid zone described in *Chapter Eight* also corresponds to a discrete elevational zone intermediate between that occupied by the two parental species. *O. acanthium* is a temperate cooler climate thistle while *O. illyricum* is generally a warmer climate coastal species. That hybrids between these species occur at the edge of these species ranges indicates that either (i) exogenous climatically driven selection has maintained species boundaries or that (ii) endogenous selection and intrinsic hybrid unfitness have reduced genetic exchange between these species. However, in *Chapter Nine*, I demonstrate that hybrids are not intrinsically unfit in many important life history traits using controlled glasshouse conditions. Therefore, it appears most likely that ecological mechanisms determine the shape of *Onopordum* hybrid zones maintaining narrow species boundaries in Europe, and broad to non-existent boundaries in Australia.

To understand the basis for a narrow hybrid zone in Europe, and either a failure of one to develop or its collapse in Australia, I undertook an analysis of the relative intrinsic fitness of hybrids in *Chapter Nine*. A decrease in hybrid fitness has explained the maintenance of hybrid zones in several situations (Arnold & Hodges 1995, Rieseberg 1997). However, since hybrids are abundant in Australia, it was possible that any

decrease in hybrid fitness was due to external selection. To perform this analysis, I explored a range of fitness components at different life-history stages in hybrid *Onopordum*, focusing on the relative intrinsic fitness without reference to environmental variation. For seed number, individual achene mass, total achene mass output, germination of achenes, seedling growth rate, and inter-species competitive ability, hybrids were intermediate between parents. Only for the proportion of unfilled achenes did hybrids display reduced vigour relative to parents. Consequently, as hybrids were not intrinsically unfit, the maintenance and collapse of hybrid zones is probably driven by the ecological factors. Such conditions of intrinsic fitness may further promote the evolution of a suite of distinct hybrid lineages during the invasion of novel ecological conditions.

References

Abad P, Pascual B, Maroto JV, Lopezgalarza S, Vicente MJ, Alagarda J (1998) RAPD analysis of cultivated and wild yellow nutsedge (*Cyperus esculentus* L.) *Weed Science* **46**, 317-321.

Abbott RJ (1992) Plant invasions, interspecific hybridization and the evolution of new plant taxa, *Trends in Ecology and Evolution* **7**, 401-405.

Allan GJ, Clark C, Rieseberg LH (1997) Distribution of parental DNA markers in *Encelia virginensis* (*Asteraceae, Heliantheae*), a diploid species of putative hybrid origin. *Plant Systematics and Evolution* **205**, 205-221.

Anderson E. (1949) *Introgressive Hybridization*. John Wiley & Sons, Inc., New York.

Andrews TS, Morrison IN, Penner GA (1998) Monitoring the spread of ACCase inhibitor resistance among wild oat (*Avena fatua*) patches using AFLP analysis. *Weed Science* **46**, 196-199.

Arnold ML, Bennett BD (1993) Natural hybridization in Louisiana irises: genetic variation and ecological determinants. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp115-139. Oxford University Press, New York.

Arnold ML, Hodges SA (1995) Are natural hybrids fit or unfit relative to their parents? *Trends in Ecology and Evolution* **10**, 67-71.

Arnold ML, Bulger MR. Burke JM. Hempel AL, Williams JH (1999) Natural hybridization: How low can you go and still beimportant? *Ecology* **80**, 371-381.

Arriola PE, Ellstrand NC (1996) Crop-to-weed gene flow in the genus Sorghum (Poaceae); spontaneous interspecific hybridization between johnsongrass, *Sorghum halpense*, and crop sorghum *S. bicolor*. *American Journal of Botany* **83** 1153-1159.

Arriola PE, Ellstrand NC (1997) Fitness of interspecific hybrids in the genus *Sorghum*: persistence of crop genes in wild populations. *Ecological Applications* **7**, 512-518.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**, 113-148.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature* **341**, 497-503.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**, 113-148.

Bayer RJ, Starr JR (1998) Tribal phylogeny of the Asteraceae based on two non-coding chloroplast sequences, the trnL intron and trnL/trnF intergenic spacer. *Annals of the Missouri Botanical Garden* **85**, 242-256.

Beismann H, Barker JHA, Karp A, Speck T (1997) AFLP analysis sheds light on distribution of two *Salix* species and their hybrid along a natural gradient. *Molecular Ecology* **6**, 989-993.

Belbin (1983) PATN: Manual for statistical analysis. CSIRO publishing.

Bennett KD (1986) The rate of spread and population increase of forest trees during the postglacial. *Phil. Trans. R. Soc. Lon.* **314**, 523-531

Berg EE, Hamrick JL (1997) Quantification of genetic diversity at allozyme loci. *Canadian Journal Forest Resources* **27**, 415-424.

Bert TM, Arnold WS (1995) An empirical test of preictions of two competing models for the maintenance and fate of hybrid zones: both models are supported in a hard-clam hybrid zone. *Evolution* **49**, 276-289.

Boecklen WJ, Howard DJ (1997) Genetic analysis of hybrid zones: numbers of markers and power of resolution *Ecology* **78**, 2611-2616.

Boudry P, Mörchen M, Saumitoulaprade P, Vernet P, Vandijk (1993). The origin and evolution of weed beets: consequences for the breeding and release of herbicide-resistant transgenic sugar beets. *Theoretical and Applied Genetics* **87**, 471-478.

Bremer B (1996) Combined and separate analyses of morphological and molecular data in the plant family Rubiaceae. *Cladistics* **12**, 21-40.

Bremer K (1994) Asteraceae: cladistics and classification. Portland, Oregon: Timber Press.

Briese DT, Espiau C, Pouchot-Lermans A (1996) Micro-evolution in the weevil genus *Larinus*: host-race formation and speciation. *Molecular Ecology* **5**, 531-545.

Briese DT, Sheppard AW, Zwölfer H, Boldt PE (1994) Structure of the phytophagous insect fauna of *Onopordum* thistles in the northern Mediterranean basin. *Biological Journal of the Linnean Society* **53**, 231-253.

Briese DT (1990) A new biological control programme against thistles of the genus *Onopordum* in Australia. In: *Proceedings 7th International Symposium Biological Control of Weeds*. (ed. ES Delfosse). pp. 6-11, Rome, Italy.

Briese DT (1996) Phylogeny: can it help us to understand host-choice by biological control agents? *Proc. IX Int. Symp. Biol. Contr. Weeds*., Moran, V.C. & Hoffman, J.H. (eds), 19-26 Jan 1996, Stellenbosch, South Africa. UCT, pp 63-70.

Briese DT (1996) Potential impact of the stem-boring weevil *Lixus cardui* on the growth and reproductive capacity of *Onopordum* thistles. *Biocontrol Science & Technology* **6**, 251-261.

Briese DT, Lane D, Hyde-Wyatt BH, Crocker J, Diver RG (1990) Distribution of thistles of the genus *Onopordum* in Australia. *Plant Protection Quarterly* **5**, 23-27.

Brochmann C (1987) Evaluation of some methods for hybrid analysis, exemplified by hybridization in *Argyranthemum* (*Asteraceae*). *Nordic Journal of Botany* **7**, 609-630.

Brown AHD, Weir BS (1983) Measuring genetic variability in plant populations. (eds.) Tanksley SD; Orton Part A TJ. In Isozymes in Plant Genetics and Breeding Elsevier Science Publ.: Amsterdam. 219-239 pp.

Brown AHD (1990) Genetic characterization of plant mating systems. (eds.) Brown HD; Clegg MT; Kahler AL; Weir S. In Plant Population Genetics, Breeding and Genetic esources Sinauer Assoc: Sunderland. 145-162 pp.

Burke JM, Carney SE, Arnold ML (1998) hybrid fitness in the Louisiana irises: analysis of parental and F1 performance. *Evolution* **52**, 37-43.

Cain ML, Andreasen V, Howard DJ (1999) Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. *Evolution* **53**, 1343-1353.

Campbell DR, Waser NM, Wolf PG (1998) Pollen transfer by natural hybrids and parental species in an *Ipomopsis* hybrid zone. *Evolution* **52**, 1602-1611.

Chaboudez P, Sheppard AW (1992) Are particular weeds more amenable to biological control? A reanalysis of mode of reproduction and life history. Proc. VIII Int. Symp. Biological Control of Weeds. pp. 95-102.

Charlesworth D (1995) Evolution under the microscope. *Current Biology* **5**, 835-836.

Comes HP, Abbott RJ (1999) Reticulate evolution in the Mediterranean species complex of *Senecio* sect. *Senecio*: Uniting phylogenetic and population level approaches. In: Hollingsworth, P.M., Bateman R.M., Gornall R.J. (eds) Molecular Systematics and Plant Evolution. Taylor & Francis Ltd.

Cummings CL, Alexander HM, Snow AA (1999) Increased pre-dispersal seed predation in sunflower crop-wild hybrids. *Oecologia* **121**, 330-338.

Darmency H (1994) The impact of hybrids between genetically modified crop plants and their related species: introgression and weediness. *Molecular Ecology* **3**, 37-40.

Dekker J (1997) Weed diversity and weed management. *Weed Science* **37**, 237-246.

Dittrich M (1977) Cynareae, systematic review. In Heywood, V.H., Harborne, J.B., Turmer, D. L. (Eds.) The biology and chemistry of the Compositae, pp. 999-1015. London, New York. Academic Press.

Donoghue MJ (1994) Progress and prospects in reconstructing plant phylogeny. *Annals of the Missouri Botanical Garden* **81**, 405-418.

Donoghue MJ, Mathews S (1998) Duplicate genes and the root of angiosperms, with an example using phytochrome sequences. *Molecular Phylogenetics & Evolution* **9**, 489-500.

Dowling TE, Moritz C, Palmer JD, Rieseberg LH (1996) Analysis of Fragments and Restriction Sites. In Hillis DM, Moritz C, Mable BK (Eds.) Molecular Systematics. pp 249-282. Sinauer. Sunderland. Massachusetts.

Doyle JJ, Morgante M, Tingey SV, Powell W (1998) Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (Glycine subgenus Glycine). *Molecular Biology & Evolution* **15**, 215-218.

Dress WJ (1966) Notes on cultivated Compositae. *Onopordum. Baileva*, **14**, 75-86.

Dumolinlapegue S, Pemonge MH, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology* **6**, 393-397.

Emms SK, Arnold ML (1997) The effect of habitat on parental and hybrid fitness: reciprocal transplant experiments with Louisiana irises. *Evolution* **51**, 1112-1119.

Emms SK, Hodges SA, Arnold ML (1996) Pollen-tube competition, siring success, and consistent asymmetric hybridization in Louisiana irises. *Evolution* **50**, 2201-2206.

Estoup A, Tailliez C, Cournet JM, Solignac M (1995) Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution* **12**, 1074-1084.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitocondrial DNA restriction sites. *Genetics* **131**, 479-491.

Fenner M (1985) *Seed Ecology*. Chapman and Hall, London.

Floate KD, Martinsen GD, Whitham TG (1997) Cottonwood hybrid zones as centres of abundance for gall aphids in western North America - importance of relative habitat size. *Journal of Animal Ecology* **66**, 179-188.

Floate KD, Whitham TG and Keim P (1994) Morphological versus genetic markers in classifying hybrid plants. *Evolution* **48**, 929-930.

Franco J. do Amaral (1976) *Onopordum*. In: *Flora Europea, Vol 4* (ed. TG Tutin), pp. 224, Cambridge University Press, Cambridge.

Friesen VL, Congdon BC, Walsh HE, Birt TP (1997) Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. *Molecular Ecology* **6**,1047-1058.

Fritz RS, Nicholsorians CM, Brunsfeld SJ (1997) Interspecific hybridization of plants and resistance to herbivores - hypotheses, genetics, and variable responses in a diverse herbivore community. *Oecologia* **97**, 106-117.

Fuertes Alguilar J, Rossello JA, Nieto Feliner G (1999) Molecular evidence for the compilospecies model of reticulate evolution in Armeria. *Systematic Biology* **48**, 735-754.

Fuytuma DJ, Keese MC, Funk DJ (1995) Genetic constraints on macroevolution: The evolution of host affiliation in the leaf beetle genus *Ophraella*. *Evolution* **49**, 797-809.

Geritz SAH (1995) Evolutionary stable seed polymorphism and small-scale spatial variation in seedling density. *American Naturalist* **146**, 685-707.

Gielly L, Taberlet P (1996) A phylogeny of the European gentians inferred from chloroplast trnL(UAA) intron sequences. *Botanical Journal of the Linnean Society* **120**, 57-75.

Gilmore S, Weston PH and Thomson JA (1992) A simple, rapid, inexpensive and widely applicable technique for purifying plant DNA. *Australian Systematic Botany* **6**, 139-148.

Goldstein DB, Pollock DD (1997) Launching microsatellites: A review of mutation processes and methods of phylogenetic inference. *Journal of Heredity* **88**, 335-342.

Gonzalez Sierra G, Perez Morales C Penas Merino A, Rivas-Martinez S (1992) Revisions taxanomica de las especies ibericas del genero *Onopordum* L. *Candollea* **47**, 181-213.

Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics* **27**, 857-874.

Groves RH, Burdon JJ, Kay PE (1990) Demography and genetics of *Onopordum* in southern New South Wales. *Journal of Ecology* **78**, 47-55.

Gupta PK, Balyan IS, Sharma PC, Ramesh B (1996) Microsatellites in plants - a new class of molecular markers. *Current Science* **70**, 45-54.

Hadrys H, Balick M, Schierwater B (1992) Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Molecular Ecology* **1**, 55-63.

Haig D, Westoby M (1988) Inclusive fitness, seed resources and maternal care. *Plant Reproductive Ecology: Patterns and Strategies* (eds Loverr Doust J, Lovett Doust L) pp. 60-79. Oxford University Press, New York.

Harper JL, Lovell PH, Moore KG (1970) The shapes and sizes of seeds. *Annual Review of Ecology and Systematics* **1**, 327-356.

Harrison RG (1993) Hybrids and hybrid zones: historical perspective. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp3-12, Oxford University Press, New York.

Hatfield T, Schluter D (1999) Ecological speciation in sticklebacks: Environment-dependent hybrid fitness. *Evolution* **53**, 866-873.

Hauser TP, Jorgensen RB, Ostergard D. (1997) Preferntial exclusion of hybrids in mixed pollinations between oilseed rape (*Brassica napus*) and weedy *B. campestris* (*Brassicaceae*). *American Journal of Botany* **84**, 756-762.

Hauser TP, Shaw RG, Ostergard H (1998) Fitness of F1 hybrids between weedy *Brassica rapa* and oilseed rape (*B. napus*). *Heredity* **81**, 429-435.

Hercus MJ, Hoffmann AA (1999) Does interspecific hybridization influence evolutionary rates? An experimental study of laboratory adaptation inhybrids between *Drosophila serrata* and *Drosophila birchii*. *Proceedings of the Royal Society of London - Series B: Biological Sciences* **266**, 2195-2200.

Heslewood MM, Elphinstone MS, Tidemann SC, Baverstock PR (1998) Myoglobin intron variation in the gouldian finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoresis. Electrophoresis **19**, 142-151.

Heun M, Helentjaris T (1993) Inheritance of RAPDs in F1 hybrids of corn. *Theoretical and Applied Genetics* **85**, 961-968.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247-276.

Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL V: improved software for multiple sequence alignment. *Computer Applications in the Biosciences* **8**,189-191.

Hillis DM, Mable BK, Larson A, Davis, SK, Zimmer EA (1996) Sequencing and Cloning. In Hillis DM, Moritz C, Mable BK (Eds.) Molecular Systematics. pp 321-347. Sinauer. Sunderland. Massachusetts.

Hollingsworth ML, Hollingsworth PM, Jenkins GI, Bailey JP (1998) The use of molecular markers to study patterns of genotypic diversity in some invasive alien *Fallopia* spp. (Polygonaceae). *Molecular Ecology* **7**, 1681-1691.

Holt RD, Hochberg ME (1997) When is biological control evolutionarily stable (or is it?) *Ecology* **78**, 1673-1683.

Hongtrakul V, Slabaugh MB, Knapp SJ (1998) DFLP, SSCP, and SSR markers for delta-9-stearoyl-acyl carrier protein desaturases strongly expressed in developing seeds of sunflower: intron lengths are polymorphic among elite inbred lines. *Molecular Breeding* **4**, 195-203.

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing Buffalograss [*Buchloe dactyloides* (Nutt.) Engelm.] *Theoretical and Applied Genetics* **86**, 927-934.

Huntley B, Birks HJB (1983) An atlas of past and present pollen maps for Europe. Cambridge: Cambridge University Press.

Huson DH (1998) SPLITSTREE, analyzing and visualizing evolutionary data. *Bioinformatics* **14**, 68-73.

Huys G, Coopman R, Janssen P, Kersters K (1996) High-resolution genotypic analysis of the genus *Aeromonas* by AFLP fingerprinting. International Journal of Systematic Bacteriology **46**, 572-580.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends in Ecology and Evolution **11**, 424-429.

Jasieniuk M, Brulebabel AL, Morrison IN (1996) The evolution and genetics of herbicide resistance in weeds, *Weed Science*, **44**, 176-193 .

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, Vandewiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, AertR, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding **3**, 381-390.

Jordan NR, Jannink JL (1997) Assessing the practical importance of weed evolution - a research agenda. *Weed Research*, **37**, 237-246.

Julien MH, Griffiths MW (1998) Biological control of weeds: A world catalogue of agents and their target weeds. 4th Edition. CABI Publ., Wallingford.

Kardoulus JP, Van Eck HJ, Van Den Berg R (1998) The potential of AFLPs in biosystematics: a first application in Solanum taxonomy (Solanaceae). *Plant Systematics and Evolution* **210**, 87-103.

Karp A, Edwards KJ (1995) Techniques for the analysis, characterization and conservation of plant genetic resources: molecular techniques in the analysis of the extent and distribution of genetic diversity. Report of an IPGRI Workshop, October 1995, Rome, Italy. IPGRI: Italy.

Katzir N, Portnoy V, Tzuri G, Castejonmunoz M, Joel DM (1996) Use of Random Amplified Polymorphic DNA (RAPD) markers in the study of the parastic weed *Orobanche. Theoretical and Applied Genetics* **93**, 367-372.

Kearsey MJ (1998) QTL analysis in plants: where are we now? Heredity **80**, 137-142.

Kelley ST, Farrell BD (1998) Is specialization a dead end? The phylogeny of host use in Dendroctonus bark beetles. *Evolution* **52**, 1731-1743.

Kelly JK, Noor MAF (1996) Speciation by reinforcement: a model derived from studies of *Drosophila. Genetics* **143**, 1485-1497.

Kim SC, Rieseberg LH (1999) Genetic architecture of species differences in annual sunflowers: Implications for adaptive trait introgression. *Genetics* **153**, 965-977.

Kopf A, Rank NE, Roinen H, Julkunen-Tiitto R, Pasteels JM, Tahlvanainen J (1998) The evolution of host-plant use and sequestration in the leaf beetle genus *Phratora* (Coleoptera: Chrysomelidae). *Evolution* **52**, 517-528.

Krauss SL (1996) A multivariate analysis of geographic variation in morphology in *Persoonia mollis* (Proteaceae). *Plant Systematics and Evolution* **202**, 65-86.

Krauss SL (1999) Complete exclusion of nonsires in an analysis of paternity in a natural plant population using amplified fragment length polymorphism (AFLP). *Molecular Ecology* **8**, 217-226.

Krauss SL, Peakall R (1998) An evaluation of the AFLP fingerprinting technique for the analysis of paternity in natural populations of *Persoonia mollis* (Proteaceae). *Australian Journal of Botany* **46**, 533-546.

Kruskall JB (1964) Non-metric multidimensional scaling: a numerical method. *Psychometrika* **29**, 115-129.

Lacy RC (1987) Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology* **1**, 143-158.

Lessa EP (1990) Multidimensional analysis of geographic genetic structure. *Systematic Zoology* **39**, 242-252.

Levin DA, Franciscortega J, Jansen RK (1996) Hybridization and the extinction of rare plant species. *Conservation Biology* **10**,10-16.

Lin J, Kuo J (1995) AFLP™: A novel PCR-based assay for plant and bacteria DNA fingerprinting. *Focus* **17**, 66-60.

Lipow SR, Wyatt R. (1999) Diallel crosses reveal patterns of variation in fruit-set, seed mass, and seed number in *Asclepias incarnata*. *Heredity* **83**, 310-318.

Louda SM, Kendall D, Connor J, Simberloff D (1997) Ecological effects of an insect introduced for the biological control of weeds. *Science* **277**, 1088-1090.

Maccallum CJ Nurnberger B Barton NH Szymura JM (1998) Habitat preference in the *Bombina* hybrid zone in Croatia. *Evolution* **52**, 227-239.

Maheswaran M, Subudhi PK, Nandi S (1997) Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. *Theoretical & Applied Genetics* **94**, 39-45.

Marohasy J (1996) Host shifts in biological weed control: real problems, semantic difficulties or poor science. *International Journal of Pest Management* **42**, 71-75.

Mason RJ, Holsinger KE, Jansen RK (1993) Biparental Inheritance of the Chloroplast Genome in *Coreopsis* (*Asteraceae*). *Journal of Heredity* **85**, 171-173.

Meerts P, Baya T, Lefebvre C (1998) Allozyme variation in the annual weed species complex *Polygonum aviculare* (*Polygonaceae*) in relation to ploidy level and colonizing ability. *Plant Systematics and Evolution* **211**, 239-256.

Michael PW (1968) Control of biennial thistle, *Onopordum*, by amitrole and five perennial grasses. *Australian Journal of Experimental Agriculture and Animal Husbandry* **8**, 332-339.

Michael PW (1996) Necessary background for studies in the taxonomy of *Onopordum* in Australia. *Plant Protection Quarterly* **11**, 239-241.

Michalakis Y, Derancourt F, Noel V, Espiau C (1993) A 1st estimation of the mating system of *Onopordum. illyricum* (L.) and its relevance to the invasion of Australia. *Acta Oecologica-International Journal of Ecology* **14**, 539-545.

Milligan BG, McMurry CK (1993) Dominant vs. codominant genetic markers in the estimation of male mating success. *Molecular Ecology* **2**, 275-283.

Mindell DP, Thacker CE (1996) Rates of molecular evolution: phylogenetic issues and applications. *Annual Review of Ecology and Systematics* **27**, 279-303.

Mitter C, Farrell B, Futuyma DJ (1991) Phylogenetic studies of insect-plant interactions: Insights into the genesis of diversity. *Trends in Ecology and Evolution* **6**, 290-293.

Moodie M, Finch RP, Marshall G (1997) Analysis of genetic variation in wild mustard (*Sinapis arvensis*) using molecular markers. *Weed Research* **45**, 102-107.

Moore WS, Price JT (1993) Nature of selection in the northern flicker hybrid zone and its implications for speciation theory. In: *Hybrid zones and the Evolutionary Process* (ed. Harrison RG), pp. 196-225. Oxford University Press, New York.

Mörchen M, Cuguen J, Michaelis G, Hanni C, Saumitouplaprade P (1996) Abundance and length polymorphism of microsatellite repeats in *Beta vulgaris* L. *Theoretical and Applied Genetics* **92**, 326-333.

Morell MK, Peakall R, AppelsR, Preston LR, Lloyd HL (1995) DNA profiling techniques for plant variety identification. *Australian Journal of Experimental Agriculture* **35**, 807-819.

Morrison SM, Scott JK (1996) Variation inpopulations of *Tribulus terrestris* (Zygophyllaceae). 3. Isozyme analysis. *Australian Journal of Botany* **44**, 201-212.

Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* **14**, 389-394.

Nissen SJ, Masters RA, Lee DJ, Rowe ML (1995) DNA-based marker systems to determine genetic diversity of weedy species and their application to biocontrol. *Weed Science* **43**, 504-513.

O'Hanlon PC, Peakall R, Briese D (1999) Amplified fragment length polymorphism (AFLP) reveals introgression in weedy *Onopordum* thistles: hybridization and invasion. *Molecular Ecology* **8**, 1259-1246.

O'Hanlon and Peakall (2000) A simple method for the detection of size homoplasy among AFLP fragments. Molecular Ecology, 9, 000-000.

O'Hanlon, Peakall and Briese (2000) A review of new PCR-based genetic markers and their utility to weed ecology. *Weed Research*, **40**, 239-254.

O'Hanlon, Briese, Peakall, (2000)[b] Know your enemy: The use of molecular biology in the *Onopordum* biological control project. Proc. X Int. Symp. Biol. Contr. Weeds, Spencer, N.R. & Nowierski, R. (eds), pp281-289. 5-9 July 1999, Bozeman, Montana, Elk River Printers, Sidney.

Okoli CAN, Shilling DG, Smith RL, Berwick TA (1997) Genetic diversity in Purple Nutsedge (*Cyperus rotundus* L.) and Yellow Nutsedge (*Cyperus esculentus* L.). *Biological Control* **8**, 111-118.

Oldroyd BP. Cornuet JM. Rowe D. Rinderer TE. Crozier RH.(1995) racial admixture of *Apis mellifera* in Tasmania, Australia: Similarities and differences with natural hybrid zones in Europe. *Heredity* **74**, 315-325.

Parker PG, Snow AA, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* **79**, 361-382.

Parris MJ. (1999) Hybridization in leopard frogs (*Rana pipiens* complex): Larval fitness components in single-genotype populations andmixtures. *Evolution* **53**, 1872-1883.

Parris MJ. Semlitsch RD. Sage RD. (1999) Experimental analysis of the evolutionary potential of hybridization in leopard frogs (Anura : Ranidae). *Journal of Evolutionary Biology* **12**, 662-671.

Parsons WT, Cutherbertson EG (1992) *Onopordum*. In: *Noxious Weeds of Australia*. Inkata Press, Melbourne.

Peakall R, Smouse P (1998) GenAlEx: Genetic Analysis in Excel v4.1. Software Support for Genetic Analysis for Population and Biogeographic Studies. A National Workshop held at the Australian National University July 2-7, 1998.

Peakall R (1997) PCR-Based genetic markers and their applications to turfgrass breeding. *International Turfgrass Society Research Journal* **8**, 243-257.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Molecular Biology & Evolution* **15**,1275-1287.

Peakall R, Smouse PE, Huff DR (1995) Evolutionary implications of allozyme and RAPD variation in diploid populations of Buffalograss. (*Buchloë dactyloides* (Nutt. Engelm.). *Molecular Ecology* **4**, 135-147.

Peirce JR (1996) The relevance of variation in thistles to herbicidal control. *Plant Protection Quarterly* **11**, 277-279.

Penner GA, Bush A, Wise R, Kim W, Domier L, Kasha K, Laroche A, Scoles G, Motnar SJ, Fedak, G (1993) Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. *PCR Methods and Applications* **2**, 341-353.

Petit DP (1997) Generic interrelationships of the *Cardueae (*Compositae): a cladistic analysis of morphological data. *Plant Systematics and Evolution* **207**, 173-203.

Pettit W, Briese DT, Walker A (1996) Aspects of thistle population dynamics with reference to *Onopordum*. *Plant Protection Quarterly* **11**, 232-235.

Podani J (1995) *SYN-TAX 5.02 Mac: Computer Programs for Multivariate Data Analysis on the Macintosh system*. Scientia Publishing, Budapest.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* **2**, 225-238.

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends in Genetics* **9**, 275-280.

Ramser J, Weising K, Lopezperalta C, Terhalle W, Terauchi R, Kahl G, (1997) Molecular marker based taxonomy and phylogeny of Guinea yam (*Dioscorea rotundata D. cayenensis*). *Genome* **40**, 903-915.

Rees M, Sheppard A, Briese D, Mangel M (1999) Evolution of size-dependent flowering in *Onopordum illyricum*: A quantitative assessment of the role of stochastic selection pressures. *American Naturalist* **154**, 628-651.

Rieseberg LH (1997) Hybrid origins of plant species . *Annual Review of Ecology & Systematics* **28**, 359-389.

Rieseberg LH, Carney SE (1998) Plant hybridization. *New Phytologist* **140**, 599-624.

Rieseberg LH, Ellstrand NC (1993) What can molecular and morphological markers tell us about plant hybridization. *Critical Reviews in Plant Sciences* **12**, 213-241.

Rieseberg LH, Linder CR (1999) Hybrid classification: Insights from genetic map-based studies of experimental hybrids *Ecology* **80**, 361-370.

Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM (1996) Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science* **272**, 741-745.

Rieseberg LH (1996) Homology among RAPD fragments in interspecific comparisons. *Molecular Ecology* **5**, 99-105.

Rieseberg LH, Baird SJE, Desrochers AM (1998) Patterns of mating in wild sunflower hybrid zones. *Evolution* **52**, 713-726

Rieseberg RH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid zones and the evolutionary process* (ed. Harrison RG), pp. 70-109, Oxford University Press, New York.

Roa AC, Maya MM, Duque MC, Tohme J, Allem AC, Bonierbale MW (1997) AFLP analysis of relationships among cassava and other manihot species. *Theoretical & Applied Genetics* **95**, 741-750.

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. *Theoretical and Applied Genetics* **90**, 43-48.

Rousset F, Raymond M (1997) Statistical analyses of population genetic data: new tools, old concepts. *Trends in Ecology and Evolution* **12**, 313-317.

Rowe ML, Lee DJ, Nissen SJ, Bowditch BM, Masters RA (1997) Genetic variation in North American leafy spurge (*Eurphorbia esula*) determined by DNA markers. *Weed Science* **45**, 446-454.

Runyeon H, Prentice HC (1996) Genetic structure in the species-pair *Silene vulgaris*, and *S. uniflora* (*Caryphyllaceae*) on the Baltic Island of Oland. *Ecogeography* **19**, 181-193.

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W, Waugh R (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical &Applied Genetics* **95**, 714-722.

Ruvolo M (1994) Molecular evolutionary processes and conflicting gene trees: the hominid case. *American Journal of Physical Anthropology* **94**, 89-113.

Saiki RK, Gelfund DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. *Science* **239**, 487-491.

Sakai S, Sakai A (1995) Flower size-dependent variation in seed size: theory and a test. *American Naturalist* **145**, 918-934.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning. Chapter 13, p 45. Cols Springs Harbor Laboratory Press.

Sang T, Crawford DJ, Stuessy TF, (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (*Paeoniaceae*). *American Journal of Botany* **84**, 1120-1136 .

Scott LJ, Lange CL, Graham GC, Yetaes DK (1998) Genetic diversity and origin of Siam Weed (*Chromolaena odorata*) in Australia. *Weed Technology* **12**, 27-31.

Sharma PC, Huttel B, Winter P, Kahl G, Gardner RC, Weising K (1995) The potential of microsatellites for hybridization and polymerase chain reaction-based DNA fingerprinting of chickpea (*Cicer arietinum*) and related species. *Electrophoresis* **16**, 1755-1761.

Sharma SK, Knox MR, Ellis THN (1996) AFLP analysis of the diversity and phylogeny of lens and its comparison with RAPD analysis. *Theoretical and Applied Genetics* **93**, 751-758.

SigmaStat (1998) SPSS. Inc.

Simberloff D, Stiling P (1996) How risky is biological control? *Ecology* **77**, 1965-1974.

Smith CC, Fretwell SD (1974) The optimal balance between size and number of offspring. *American Naturalist* **108**, 499-506.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual mutiallele and multilocus genetic structure. *Heredity* **82**, 561-573.

Soltis PS, Soltis DE, Doyle JJ (1992) Molecular systematics of plants. New York, London: Chapman and Hall.

Stammers M, Harris J, Evans GM, Hayward MD, Forster JW (1995) Use of random PCR (RAPD) technology to analyse phylogenetic relationships in the *Lolium festuca* complex. *Heredity* **74**, 19-27.

Strefeler MS, Darmo E, Becker RL, Katovich EJ (1996) Isozyme variation in cultivars of purple loosestrife (*Lythrum* sp.). *Hortscience* **31**, 279-282.

Sun M (1997) Population genetic structure of yellow starthistle (*Centaurea solstitialis*), a colonizing weed in the western United States. *Canadian Journal of Botany* **75**, 1470-1478.

Susanna A, Jacas NG, Soltis DE, Soltis PS (1995) Phylogenetic relationships in the tribe Cardueae (Asteraceae) based on ITS sequences. *American Journal of Botany* **82**, 1056-1068.

Susko DJ, Lovett-Doust L (2000) Patterns of seed mass variation and their effects on seedling traits in *Alliaria petiolata* (Brassicaceae). *American Journal of Botany* **87**, 56-66.

Swofford (1999). PAUP 4.0b. Phylogenetic Analysis Using Parsimony - for Wundows 95. Sinauer Associates, Inc. Sunderland, Massachusetts.

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic Inference: in Molecular Systematics (ed) Hillis DM, Moritz C, Mable BK. Massachusetts.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**, 453-464.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**, 1105-1109.

Telenius A, Torstensson P (1999). Seed type and seed size variation in the heteromorphic saltmarsh annual *Spergularia salina* along the coast of Sweden. *Plant Biology* **1**, 585-593.

Temme DH (1986) Seed size variability: a consequence of variable genetic quality among offspring? *Evolution* **40**, 414-417.

Templeton AR (1983) Convergent evolution and non-parametric tests from restriction fragment and DNA sequence data. Pp151-179. In Weir B (ed.) Statistical analysis of DNA sequence data. Marcel Dekker, New York.

Thompson JD (1999) Population differentiation in Mediterranean plants: insights into colonization history and the evolution and conservation of endemic species. *Heredity* **82**, 229-236.

Thompson JN (1984) Variation among individual seed masses in *Lomatium grayi* (Umbelliferae) under controlled conditions: magnitude and partitioning the variance. *Ecology* **65**, 626-631.

Thormann CE, Ferreira ME, Camargo LEA, Tivang JG, Osborn TC (1994) Comparison of RFLP and RAPD markers in estimating genetic relationships within and among cruciferous species. *Theoretical and Applied Genetics* **88**, 973-980.

Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA (1972) Flora Europea, Vol 4 , Cambridge University Press, Cambridge.

Upholt WB (1977) Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. *Nucleic Acids Research* **4**, 1257-1265.

Vanham RC, Hart H, Mes THM, Sandbrink JM (1994) Molecular evolution of non-coding regions of chloroplast genome in the *Crassulaceae* and related species. *Current Genetics* **25**, 558-566.

Vaughton G, Ramsey M (1997) Seed mass variation in the shrub *Banksia spinulosa* (*Proteaceae*): Resource constraints and pollen source effects. *International Journal of Plant Sciences* **158**, 424-431.

Vaughton G, Ramsey M (1998) Sources and consequences of seed mass variation in *Banksia marginata* (*Proteaceae*). *Journal of Ecology* **86**, 563-573.

Venable DL, Brown JS (1988) The selective interactions of dispersal, dormancy and seed size as adaptations for reducing risk in variable environments. *American Naturalist* **131**, 360-384.

Vila M, Dantonio CM (1998) Fitness of invasive *Carprobrotus* (*Azioaceae*) hybrids in coastal California. *Ecoscience* **5**, 191-199.

Vos P, Hodgers R, Bleeker M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407-4414.

Wagenitz G (1997) The impact of molecular methods on the systematics of angiosperms. *Bot. Act.* **110**, 274-281.

Wagner WH Jr (1970) Biosystematics and evolutionary noise. *Taxon* **19**, 146-151.

Wajnberg E, Scott JK, Quimby PC (2000) Evaluating Indirect Ecological Effects of Biological Control. IOBC, Montpellier.

Wapshere AJ (1974) A strategy for evaluating the safety of organisms for biological weed control. *Annals of Applied Biology*, **77**, 200-221.

Wang H, Mcarthur ED, Sanderson SC, Graham JH, Freeman DC (1997) narrow hybrid zone between two subspecies of Big Sagebrush (*Artemisia tridentata*, *Asteraceae*) .4. Reciprocal transplant experiments. *Evolution* **51**, 95-102.

Wang H. McArthur ED. Freeman DC. (1999) Narrow hybrid zone between two subspecies of big sagebrush (*Artemisia tridentata*: *Asteraceae*). IX. Elemental uptake and niche separation. *American Journal of Botany* **86**, 1099-1107.

Wang RL, Wendel JF, Dekker JH (1995) Weedy adaptation in *Setaria* spp. 1. Isozyme analysis of genetic diversity and population genetic structure in *Setaria virdis*. *American Journal of Botany* **82**, 308-317.

Warwick SL, Bain JF, Wheatcroft R, Thompson BK (1989) Hybridization and introgression in *Carduus nutans* and *C. acanthoides* re-examined. *Systematic Botany* **14**, 476-494.

Waser NM (1993) Population structure, optimal outbreeding, and asortative mating in angiosperms. *The natural history of inbreeding and outbreeding: theoretical and empirical perspectives* (ed Thornhill NW) pp 173-199. University of Chicago Press, Chicago.

Weber JL (1990) Human DNA polymorphisms and methods of analysis. *Current Opinion in Biotechnology* **1**, 166-171.

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* **18**, 7213-7218.

Whitham TG, Morrow PA, Potts BM (1994) Plant hybrid zones as centers of biodiversity - the herbivore community of 2 endemic Tasmanian eucalypts. *Oecologia* **97**, 481-490.

Whitkus R, Doebley J, Wendel JF (1994) Nuclear DNA markers in systematics and evolution. In Phillips RL, Vasil IK (Eds.): DNA-based markers in plants, pp118-141. Dordrecht: Kluwer.

Wiklund A (1992) The genus *Cynara* L. (*Asteraceae*-Cardueae). *Botanical Journal of the Linnean Society* **109**, 75-123.

Wilhelm R, Hilbish TJ (1998) Assessment of natural selection in a hybrid population of Mussels; Evaluation of exogenous vs. endogenous selection models. *Marine Biology* **131**, 505-514.

Williams JGK, Kubelik AR, Livak KJ, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**, 6531-6535.

Withers TM, Barton Browne L, Stanley J (1999) Host Specificity Testing in Australasia: Towards Improved Assays for Biological Control. Scientific Publ., Indooroopilly.

Witsenboer H, Vogel J, Michelmore RW (1997) Identification, genetic localization, and allelic diversity of Selectively Amplified Microsatellite Polymorphic Loci in lettuce and wild relatives (*Lactuta* spp.). *Genome* **40**, 923-936.

Wolfe L (1995) The genetics and ecology of seed size variation in a biennial plant, *Hydrophyllum appendiculatum* (Hydrophyllaceae). *Oecologia* **101**, 343-352.

Wu K, Jones R, Danneberger L, Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research* **22**, 3257-3258.

Zhang JH, Hamill AS. (1997) Seed weight, intraspecific competition, and plant performance in *Abutilon theophastri*. *Canadian Journal of Botany-Revue Canadienne de Botanique* **75**, 1614-1620.

Zhang JH. (1998). Variation and allometry of seed weight in *Aeschynomene americana*. *Annals of Botany* **82**, 843-847.

Zietkiewicz E, Raflaski A, Labuda,D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* **20**, 176-183.